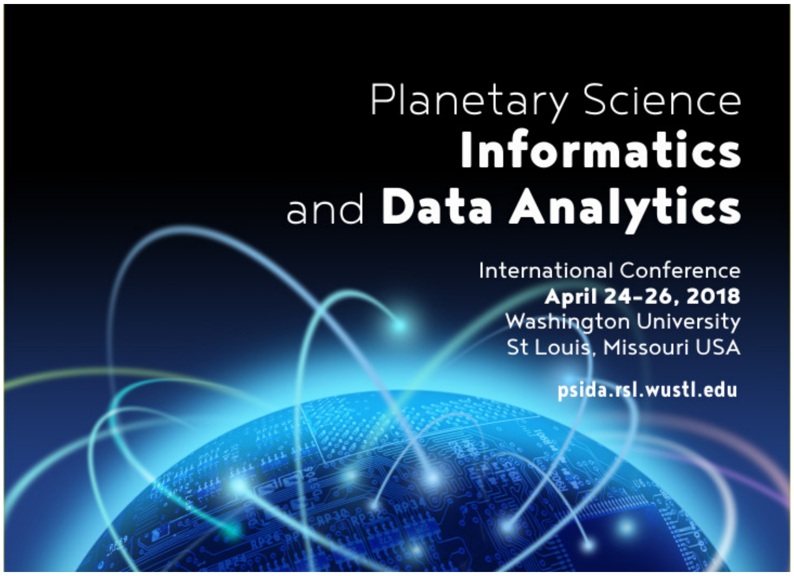


Program

Lunar and Planetary Institute   3600 Bay Area Boulevard   Houston TX 77058-1113

Planetary Science Informatics and Data Analytics Conference

April 24–26, 2018 • St. Louis, Missouri

Institutional Support

NASA Planetary Data System Geosciences

Lunar and Planetary Institute

Chairs

Tom Stein

Washington University, St. Louis, USA

Dan Crichton

Jet Propulsion Laboratory, Pasadena, USA

Program Committee

Alphan Altinok
Jet Propulsion Laboratory, Pasadena, USA
Christophe Arviset
European Space Agency, Madrid, Spain
Oleg Batanov
Russian Space Research Institute, Moscow, Russia
Baptiste Cecconi
Paris Observatory, Paris, France
Karin Eichentopf
German Aerospace Center, Berlin, Germany
Ashish Mahabal
California Institute of Technology, Pasadena, USA
Ana-Catalina Plesa
German Aerospace Center, Berlin, Germany
Jody O'Sullivan
Washington University, St. Louis, USA
Kiri Wagstaff
Jet Propulsion Laboratory, Pasadena, USA

Abstracts for this workshop are available via the workshop website at
https://psida.rsl.wustl.edu/
Abstracts can be cited as
Author A. B. and Author C. D. (2018) Title of abstract. In *Planetary Science Informatics and Data Analytics Conference,* Abstract #XXXX. LPI Contribution No. 2082, Lunar and Planetary Institute, Houston.

Guide to Sessions

Tuesday, April 24, 2018

9:00 a.m.	Knight Center Room 200	Keynote Session I
10:45 a.m.	Knight Center Room 200	Data Architectures, Management, and Data Technologies
3:40 p.m.	Knight Center Room 200	Scalable Data Processing

Wednesday, April 25, 2018

9:00 a.m.	Knight Center Room 200	Data Archiving
11:15 a.m.	Knight Center Room 200	Keynote Session II
1:30 p.m.	Knight Center Room 200	Visualization
5:00–7:00 p.m.	Knight Center Room 351	Poster Session

Thursday, April 26, 2018

9:00 a.m.	Knight Center Room 200	Science Applications
1:30 p.m.	Knight Center Room 200	Data-Driven Discovery and Analytics

Program

Tuesday, April 24, 2018
KEYNOTE SESSION I
9:00 a.m. Knight Center Room 200

9:00 a.m. Conference Opening and Welcome

9:15 a.m. Larry James *
Keynote Address
Detailed information related to this keynote address will be available soon.

Larry James, Deputy Director, Jet Propulsion Laboratory, is responsible for the day-to-day management of the JPL's resources and activities, managing the Laboratory's solar system exploration, Mars, astronomy, physics, Earth science, interplanetary network programs, and all business operations.

10:15 a.m. BREAK

Tuesday, April 24, 2018
DATA ARCHITECTURES, MANAGEMENT, AND DATA TECHNOLOGIES
10:45 a.m. Knight Center Room 200

Scalable data management and database systems, planetary information models and semantic technologies, search architectures and technologies, cloud-based architectures, application of open source data technologies.

Chairs: **Oleg Batanov**
Baptiste Cecconi

10:45 a.m. Arviset C. * Barbarisi I. Besse S. Barthelemy M. de Marchi G. Docasal R. Fraga D. Grotheer E. Heather D. Laantee C. Lim T. Macfarlane A. Martinez S. Montero A. Osinde J. Rios C. Saiz J. Vallat C.
ESA Planetary Science Archive Architecture and Data Management **[#6003]**
The Planetary Science Archive is the European Space Agency repository of science data from all planetary science and exploration missions. This paper presents PSA's content, architecture, user interfaces, and the relation between the PSA and IPDA.

11:05 a.m. Costa M. *
SPICE for ESA Planetary Missions **[#6008]**
The ESA SPICE Service leads the SPICE operations for ESA missions and is responsible for the generation of the SPICE Kernel Dataset for ESA missions. This contribution will describe the status of these datasets and outline the future developments.

11:25 a.m. Erard S. * Cecconi B. Le Sidaner P. Rossi A. P. Capria M. T. Schmitt B. André N. Vandaele A. C. Scherf M. Hueso R. Maöaöttaönen A. Carry B. Achilleos N. Marmo C. Santolik O. Benson K. Fernique P.
VESPA: Enlarging the Virtual Observatory to Planetary Science **[#6033]**
VESPA is an extension of the Virtual Observatory to handle planetary science data. This is a contributive system where data services can be provided by the community.

11:45 a.m. Politi R. * Capaccioni F. Giardino M. Fonte S. Capria M. T. Turrini D. De Sanctis M. C. Piccioni G.
SeaBIRD: A Flexible and Intuitive Planetary Datamining Infrastructure **[#6009]**
Description of SeaBIRD (Searchable and Browsable Infrastructure for Repository of Data), a software and hardware infrastructure for multi-mission planetary datamining, with web-based GUI and API set for the integration in users' software.

12:05 p.m. Hughes J. S. * Crichton D. J. Algermissen S. S. Cayanan M. D. Joyner R. S. Hardman S. H. Padams J. H.
Model-Driven Development for PDS4 Software and Services **[#6019]**
PDS4 data product labels provide the information necessary for processing the referenced digital object. However, significantly more information is available in the PDS4 Information Model. This additional information is made available for use, by both software and services, to configure, promote resiliency, and improve interoperability.

12:25 p.m. LUNCH

1:30 p.m. Beyer R. A. * Hare T. Radebaugh J.
The Need for a Planetary Spatial Data Clearinghouse **[#6067]**
Making the case for a need to have a planetary spatial data clearinghouse. Such a clearinghouse does not archive or house data, but acts as a catalog of spatial data hosted externally.

1:50 p.m. Laura J. * Arvidson R. E. Gaddis L. R.
The Relationship Between Planetary Spatial Data Infrastructure and the Planetary Data System **[#6005]**
This abstract describes the relationship between Planetary Spatial Data Infrastructure (PSDI) and the Planetary Data System (PDS).

2:10 p.m. Keszthelyi L. * Hagerty J. Akins S. Archinal B. Bailen M. Bland M. Edmundson K. Fergaons R. Hare T. Hayward R. Hunter M. Laura J. Sides S. Velasco M.
Update on the NASA-USGS Planetary Spatial Data Infrastructure Inter-Agency Agreement **[#6054]**
We provide an update on the NASA-USGS PSDI-IAA which provides much of the infrastructure for planetary science informatics and data analytics.

2:30 p.m. Lehnert K. * Ji P. Cai M. Evans C. Zeigler R.
MoonDB — A Data System for Analytical Data of Lunar Samples **[#6062]**
MoonDB is a data system that makes analytical data from the Apollo lunar sample collection and lunar meteorites accessible by synthesizing published and unpublished datasets in a relational database with an online search interface.

2:50 p.m. Stone N. * Lafuente B. Bristow T. Pires A. Keller R. M. Downs R. T. Blake D. Dateo C. E. Fonda M.
Demonstrating the Open Data Repository's Data Publisher: The CheMin Database **[#6055]**
The Open Data Repository's Data Publisher aims to provide an easy-to-use software tool that will allow researchers to create and publish database templates and related data. The CheMin Database developed using this framework is shown as an example.

3:10 p.m. BREAK

Tuesday, April 24, 2018
SCALABLE DATA PROCESSING
3:40 p.m. Knight Center Room 200

Software frameworks and systems for processing planetary data, automated workflow systems, integration of HPC, and cloud computing for data processing.

Chairs: **Christophe Arviset**
Dan Chrichton

3:40 p.m. Plesa A.-C. * Maurice M. Padovan S. Tosi N. Breuer D.
Large-Scale Numerical Simulations of Planetary Interiors **[#6023]**
We present high-performance computational methods that are employed in the mantle convection code Gaia to investigate the interior evolution of terrestrial planets.

4:00 p.m. Paris K. N. * Estes N. M. Cisneros E. Robinson M. S.
Scalable Data Processing with the LROC Processing Pipelines **[#6059]**
The LROC Science Operations Center generates and receives a lot of files of various types. We have created a way to process all of the files in a way that allows us to accommodate new file types, new processes, and new missions.

4:20 p.m. Zorzano M.-P. Martín-Torres J. Mathanlal T. Vakkada Ramachandran A. Ramirez-Luque J.-A. *
PACKMAN-Net: A Distributed, Open-Access, and Scalable Network of User-Friendly Space Weather Stations **[#6012]**
The purpose of this work is to demonstrate the operability of a network of small-sized detectors of the PACKMAN instrument, operated simultaneously to provide real time cosmic ray and solar activity monitoring over the entire planet.

Wednesday, April 25, 2018
DATA ARCHIVING
9:00 a.m. Knight Center Room 200

PDS4 data standards, tools for data preparation and ingestion, search and interoperability across planetary data archives, planetary data archives in the cloud, online data distribution services.

Chairs: **Christophe Arviset**
Dan Crichton

9:00 a.m. Announcements

9:05 a.m. Nass A. D'Amore M. *
Archive, Access, and Supply of Scientifically Derived Data: A Data Model for Multi-Parameterized Querying Where Spectral Data Base Meets GIS-Based Mapping Archive **[#6021]**
An archiving structure and reference level of derived and already published data supports the scientific community significantly by a constant rise of knowledge and understanding based on recent discussions within Information Science and Management.

9:25 a.m. Hare T. M. * Gaddis L. R.
Initial PDS4 Support for the Geospatial Data Abstraction Library (GDAL) **[#6038]**
We introduce initial support for PDS4 within the Geospatial Data Abstraction Library (GDAL). Both highlights and limitations are presented, as well as a short discussion on methods for supporting a GDAL-based workflow for PDS4 conversions.

9:45 a.m. Lim T. L. * Martinez S. Coia D. Barbarisi I. Barthelemy M. Besse S. Fraga Agudo D. Grotheer E. Heather D. Vallat C.
PDS4 Data Within the PSA — A Cross-Mission and Cross-Discipline Approach to a PDS4 Archive **[#6027]**
The cross-mission and cross-discipline nature of the European Space Agency planetary science archive has led to a degree of standardization of PDS4 data structures and attributes housed by the archive demonstrating that the PDS core could adopt a similar approach.

10:05 a.m. De Cesare C. M. * Padams J. H.
Bridging Archival Standards: Building Software to Translate Metadata Between PDS3 and PDS4 **[#6058]**
Transitioning datasets from PDS3 to PDS4 requires manual and detail-oriented work. To increase efficiency and reduce human error, we've built the Label Mapping Tool, which compares a PDS3 label to a PDS4 label template and outputs mappings between the two.

10:25 a.m. Raugh A. C. * Hughes J. S.
The PDS4 LDDTool — Information Modeling for Data Preparers **[#6017]**
We present the key features of the PDS LDDTool software tool, used by data preparers to define metadata consistent with the core PDS4 information model.

10:45 a.m. BREAK

Wednesday, April 25, 2018
KEYNOTE SESSION II
11:15 a.m. Knight Center Room 200

11:15 a.m. George Djorgovski *
Keynote Address
Detailed information related to this keynote address will be available soon.

George Djorgovski, Professor of Astronomy, Center for Data Driven Discovery Director, California Institute of Technology, is a pioneer in the field of Astroinformatics, focusing on the development of computational, mostly data-driven science. His research looks at ways in which computational and information technology changes the ways we study and understand the world around us.

12:20 p.m. LUNCH

Wednesday, April 25, 2018
VISUALIZATION
1:30 p.m. Knight Center Room 200

Planetary GIS technologies, tools for image analysis of different planetary bodies, planetary applications of VR technologies.

Chairs: **Ashish Mahabal**
Ana Plesa

1:30 p.m. Latorella K. Tisdale M. Aaron L. * Leinenveber J. Wilson J. Werynski A. Gant P.
A GIS Layer Depicting Proposed Human Landing Sites and Exploration Zones on Mars and Tool to Investigate These **[#6018]**
Mars Human Landing Site Study proposed Exploration Zones in one GIS layer, with characterizing data; and a visualization and collaboration tool for analyzing and discussing these.

1:50 p.m. DeWolfe A. W. * Larsen K. Brain D.
3D Visualization for Planetary Missions **[#6006]**
We have developed visualization tools for viewing planetary orbiters and science data in 3D for both Earth and Mars, using the Cesium Javascript library, allowing viewers to visualize the position and orientation of spacecraft and science data.

2:10 p.m. Barnouin O. S. Ernst C. M. Daly R. T. *
The Small Body Mapping Tool (SBMT) for Accessing, Visualizing, and Analyzing Spacecraft Data in Three Dimensions **[#6043]**
The free, publicly available Small Body Mapping Tool (SBMT) developed at the Johns Hopkins University Applied Physics Laboratory is a powerful, easy-to-use tool for accessing and analyzing data from small bodies.

2:30 p.m. Law E. S. * Solar System Treks Team
Planetary Surface Visualization and Analytics **[#6063]**
An introduction and update of the Solar System Treks Project which provides a suite of interactive visualization and analysis tools to enable users (engineers, scientists, public) to access large amounts of mapped planetary data products.

2:50 p.m. Poster Blitz I

3:10 p.m. BREAK

3:40 p.m. Poster Blitz II

4:00 p.m. Kidd J. N. Jr. * Selznick S. Hergenrother C. W.
Providing Observation Context via Kernel Visualization and Informatics for Planning and Data Analysis **[#6065]**
From our lessons learned and SPICE expertise, we lay out the features and capabilities of a new web-based tool to provide an accessible platform to obtain context and informatics from a planetary mission's SPICE kernels.

4:20 p.m. Heyer T. * Hiesinger H. Reiss D. Erkeling G. Bernhardt H. Lüsebrink D. Jaumann R.
The Multi-Temporal Database of Planetary Image Data: A Tool to Study Dynamic Mars **[#6013]**
The Multi-temporal Database of Planetary Image Data (MUTED) is a comprehensive tool to support the identification of surface changes and processes on Mars.

4:40 p.m. Miller M. J. * Graff T. Young K. Coan D. Whelley P. Richardson J. Knudson C. Bleacher J. Garry W. B. Delgado F. Noyes M. Valle P. Buffington J. Abercromby A.
Scientific Hybrid Reality Environments (SHyRE): Bringing Field Work into the Laboratory **[#6031]**
The SHyRE program aims to develop a scientifically-robust analog environment using a new and innovative hybrid reality setting that enables frequent operational testing and rapid protocol development for future planetary exploration.

Wednesday, April 25, 2018
POSTER SESSION
5:00–7:00 p.m. Knight Center Room 351

Trompet L. Vandaele A. C. Thomas I. R.
Tools to Manage and Access the NOMAD Data **[#6007]**
The NOMAD instrument on-board the ExoMars spacecraft will generate a large amount of data of the atmosphere of Mars. The Planetary Aeronomy Division at IASB is willing to make their tools and these data available to the whole planetary science community.

Raugh A. C. Hughes J. S.
The PDS4 Metadata Management System **[#6016]**
We present the key features of the Planetary Data System (PDS) PDS4 Information Model as an extendable metadata management system for planetary metadata related to data structure, analysis/interpretation, and provenance.

Cecconi B. Le Sidaner P. Savalle R. Bonnin X. Zarka P. Louis C. Coffre A. Lamy L. Denis L. Griessmeier J.-M. Faden J. Piker C. André N. Génot V. Erard S. King T. A. Mafi J. N. Sharlow M. Sky J. Demleitner M.
MASER: A Tool Box for Solar System Low Frequency Radio Astronomy **[#6029]**
MASER (Measuring, Analysing, and Simulating Radio Emissions) is a toolbox for solar system radio astronomy. It provides tools for reading, displaying, finding, and modeling low frequency radio datasets.

Million C.
Put Everything in a Database: Extremely Low Density Data Storage and Massive Relationality **[#6041]**
Remote sensing data might be represented far more flexibly and usefully within relational database architectures than traditional file formats.

Hardman S. Cayanan M. Hughes J. S. Joyner R. Crichton D. Law E.
A Registry for Planetary Data Tools and Services **[#6049]**
The PDS Engineering Node has upgraded a prototype Tool Registry developed by the International Planetary Data Alliance to increase the visibility and enhance functionality along with incorporating the registered tools into PDS data search results.

Aye K.-M.
Successful Design Patterns in the Day-to-Day Work with Planetary Mission Data **[#6068]**
I will describe successful data storage, data access, and data processing techniques, like embarrassingly parallel processing, that I have established over the years working with large datasets in the planetary science domain; using Jupyter notebooks.

Costa M. Grass M.
SPICE-Based Python Packages for ESA Solar System Exploration Mission's Geometry Exploitation **[#6010]**
This contribution outlines three Python packages to provide an enhanced and extended usage of SPICE Toolkit APIS providing higher-level functions and data quick-look capabilities focused on European Space Agency solar system exploration missions.

Zeng X. G. Liu J. J. Zuo W. Chen W. L. Liu Y. X.
Lunar Circular Structure Classification from Chang 'e 2 High Resolution Lunar Images with Convolutional Neural Network **[#6022]**
Circular structures are widely distributed around the lunar surface. The most typical of them could be lunar impact crater, lunar dome, et.al. In this approach, we are trying to use the Convolutional Neural Network to classify the lunar circular structures from the lunar images.

Estes N. M. Bowley K. S. Paris K. N. Silva V. H. Robinson M. S.
Taming Pipelines, Users, and High Performance Computing with Rector **[#6039]**
Rector is a high-performance job management system created by the LROC SOC team to enable processing of thousands of observations and ancillary data products as well as ad-hoc user jobs across a 634 CPU core processing cluster.

Verma R. V.
Archive Inventory Management System (AIMS) — A Fast, Metrics Gathering Framework for Validating and Gaining Insight from Large File-Based Data Archives **[#6056]**
The Archive Inventory Management System (AIMS) is a software package for understanding the distribution, characteristics, integrity, and nuances of files and directories in large file-based data archives on a continuous basis.

Giardino M.   Fonte S.   Politi R.   Ivanovski S.   Longobardo A.   Capria M. T.   Erard S.   De Sanctis M. C.
A Virtual Observatory Approach to Planetary Data for Vesta and Ceres **[#6011]**
A virtual observatory service for DAWN/VIR spectral dataset is presented, based upon the IVOA standards adapted to the planetary field. Advantages of such an approach will be discussed, especially concerning interoperability and availability.

Marmo C.   Hare T. M.   Erard S.   Cecconi B.   Minin M.   Rossi A. P.   Costard F.   Schmidt F.
FITS and PDS4: Planetary Surface Data Interoperability Made Easier **[#6024]**
This abstract describes how Flexible Image Transport System (FITS) can be used in planetary surface investigations, and how its metadata can easily be inserted in the PDS4 metadata distribution model.

Saiz J.   Barbarisi I.   Docasal R.   Rios C.   Montero A.   Macfarlane A.   Laantee C.   Besse S.   Vallat C.   Marcos J.   Arenas J.   Osinde J.   Arviset C.
PDS4 Challenges in the PSA **[#6025]**
The Planetary Science Archive (PSA) stores products from all planetary ESA missions. Adopting PDS4 as the standard for new missions, while being compatible with existing PDS3 products, has driven a design with several difficulties to overcome.

Yamamoto Y.   Ishihara Y.   Murakami S.
Data Archives and Standards for Japanese Planetary Missions **[#6026]**
JAXA prepares planetary data archives in the Planetary Data System. We discuss the current status and difficulties to develop data archives to follow standards.

Million C.   Brazier A.   King T.   Hayes A.
Announcing a Community Effort to Create an Information Model for Research Software Archives **[#6042]**
An effort has started to create recommendations and standards for the archiving of planetary science research software. The primary goal is to define an information model that is consistent with OAIS standards.

Padams J.   Cayanan M.   Hardman S.
PDS4: Harnessing the Power of Generate and Apache Velocity **[#6044]**
The PDS4 Generate Tool is a Java-based command-line tool developed by the Cartography and Imaging Sciences Nodes (PDSIMG) for generating PDS4 XML labels, from Apache Velocity templates and input metadata.

Kingston C.   Palmer E.   Stone J.   Drum M.   Neese C.   Mueller B.
Demonstration of Updated OLAF Capabilities and Technologies **[#6046]**
Ongoing improvements to the On-Line Archiving Facility (OLAF) NASA PDS tool are leading to improved usability and functionality by integration of JavaScript web application frameworks.

Zinzi A.   Longobardo A.   Giardino M.   Ivanovski S.   Capria M. T.   Palomba E.
MATISSE 2.0: New Ideas to Support Planetary Sciences **[#6002]**
The next version of MATISSE will be a brand-new one, so that, without wiping out the successful modular structure of the current version, a number of newly introduced solutions can allow overcoming existing limitations.

Estes N. M.   Silva V. H.   Bowley K. S.   Lanjewar K. K.   Robinson M. S.
Standards-Based Open-Source Planetary Map Server: Lunaserv **[#6035]**
Lunaserv is a planetary capable Web Map Service developed by the LROC SOC. It enables researchers to serve their own planetary data to a wide variety of GIS clients without any additional processing or download steps.

Grimes K. M. II   Padams J. H.   Stanboli A.   Wagstaff K. L.
Browsing the PDS Image Archive with the Imaging Atlas and Apache Solr **[#6048]**
The PDS Image Archive is home to tens of millions of images, nearly 30 million of which are associated with rich metadata. By leveraging the Solr indexing technology and the Imaging Atlas interactive frontend, we enable intuitive archive browsing.

Heyd R. S.   McArthur G. A.   Leis R.   Fennema A.   Wolf N.   Schaller C. J.   Sutton S.   Plassmann J.   Forrester T.   Fine K.
Applying the Lessons Learned from the HiRISE Ground Data System to the Development of a Modernized GDS for CaSSIS **[#6066]**
The HiRISE ground data system is a mature data processing system in operation for over 12 years. The experience gained from this system will be applied to developing a new and more modern GDS to process data from the CaSSIS instrument.

Norman C. J.   Paxman J.   Benedix G. K.   Tan T.   Bland P. A.   Towner M.
Automated Detection of Craters in Martian Satellite Imagery Using Convolutional Neural Networks **[#6004]**
Crater counting is used in determining surface age of planets. We propose improvements to martian Crater Detection Algorithms by implementing an end-to-end detection approach with the possibility of scaling the algorithm planet-wide.

Varatharajan I.   D'Amore M.   Maturilli A.   Helbert J.   Hiesinger H.
Machine Learning Approach to Deconvolution of Thermal Infrared (TIR) Spectrum of Mercury Supporting MERTIS Onboard ESA/JAXA BepiColombo **[#6015]**
Machine learning approach to spectral unmixing of emissivity spectra of Mercury is carried out using endmember spectral library measured at simulated daytime surface conditions of Mercury. Study supports MERTIS payload onboard ESA/JAXA BepiColombo.

Wagner R. V.   Henriksen M. R.   Manheim M. R.   Robinson M. S.
Using Agisoft Photoscan to Compare Terrestrial and Planetary Volcanic Features **[#6050]**
We used Agisoft Photoscan to create three high-resolution digital terrain models (DTMs) of terrestrial volcanic features. We explore the potential for using these DTMs to better understand analogous features on the Moon and on Mars.

Thursday, April 26, 2018
SCIENCE APPLICATIONS
9:00 a.m. Knight Center Room 200

Application of data technologies to planetary missions and science investigations supporting data processing, archiving and analysis, and applications to ground systems including science data centers.

Chairs: **Karin Eichentopf**
Jody O'Sullivan

9:00 a.m. Announcements

9:05 a.m. Stark A. Matz K.-D. Roatsch T. *
Multi-Mission Laser Altimeter Data Processing and Co-Registration of Image and Laser Data at DLR **[#6014]**
We designed a system for the processing and storage of large laser altimeter data sets for various past and operating laser altimeter instruments. Furthermore, we developed a technique to accurately co-register multi-mission laser and image data.

9:25 a.m. Fergason R. L. * Weller L.
The Importance of Geodetically Controlled Data Sets: THEMIS Controlled Mosaics of Mars, a Case Study **[#6030]**
Accurate image registration is necessary to answer questions that are key to addressing fundamental questions about our universe. To provide such a foundational product for Mars, we have geodetically controlled and mosaicked THEMIS IR images.

9:45 a.m. Archinal B. * IAU Working Group on Cartographic & Rotat Elements
Planetary Coordinates Recommendations from the IAU Working Group on Cartographic Coordinates and Rotational Elements **[#6047]**
We describe the work of the International Astronomical Union (IAU) Working Group on Cartographic Coordinates and Rotational Elements, our new report, and the PSDI standards support we provide for the planetary community, including for planetary informatics and data analytics.

10:05 a.m. Showalter M. R. * Ballard L. French R. S. Gordon M. K. Tiscareno M. S.
Developments in Geometric Metadata and Tools at the PDS Ring-Moon Systems Node **[#6053]**
Object-Oriented Python/SPICE (OOPS) is an overlay on the SPICE toolkit that vastly simplifies and speeds up geometry calculations for planetary data products. This toolkit is the basis for much of the development at the PDS Ring-Moon Systems Node.

10:25 a.m. BREAK

10:55 a.m. Anderson R. B. * Finch N. Clegg S. M. Graff T. Morris R. V. Laura J.
The Python Spectral Analysis Tool (PySAT) for Powerful, Flexible, and Easy Preprocessing and Machine Learning with Point Spectral Data **[#6045]**
The PySAT point spectra tool provides a flexible graphical interface, enabling scientists to apply a wide variety of preprocessing and machine learning methods to point spectral data, with an emphasis on multivariate regression.

11:15 a.m. Hare T. M. * Laura J. R.
A Sandbox Environment for the CSM Standard and SPICE **[#6040]**
We present ongoing work USGS is undertaking to provide a programming environment for the Camera Sensor Model (CSM) standard and associated SPICE information. This allows for instrument testing and experimentation outside a given production area.

11:35 a.m. Meyer B. S. *
NucleusHUB.org: A Platform for Collaboration Among Astronomers, Nuclear Astrophysicists, and Planetary Scientists **[#6057]**
The author and collaborators are developing nucleusHUB.org, built with HUBzero technology, to facilitate interaction among astronomers, nuclear astrophysicists, and planetary scientists. The site allows users to collaborate and publish online tools.

11:55 a.m. Mahanti P. * Brown H. Robinson M. S. Boyd A. Humm D. Awumah A.
Characterizing the Spectral Radiance of Lunar Permanently Shadowed Regions **[#6020]**
Lunar Reconnaissance Orbiter Narrow Angle Camera images are used to characterize reflected illumination within permanently shadowed regions at the lunar poles.

12:15 p.m. Thompson T. J. * Mahanti P.
Analysis of Simulated Temporal Illumination at the Lunar PSRs **[#6037]**
Illumination on the Moon is modeled temporally for permanently shadowed regions to lighting trends. Crater topography is used to generate viewfactor maps, which show which areas contribute most to scattered light into the primary shadows.

12:35 p.m. LUNCH

Thursday, April 26, 2018
DATA-DRIVEN DISCOVERY AND ANALYTICS
1:30 p.m. Knight Center Room 200

Methods for feature extraction and classification from planetary data, integration of search and data analytics, automated metadata tagging, ground-based and on-board data triage, data fusion, analytics on data pipelines for computation, data discovery, event detection, and reduction.

Chairs: **Alphan Altinok**
Kiri Wagstaff

1:30 p.m. Zeigler R. A. * Blumenfeld E. H. Srinivasan P. McCubbin F. M. Evans C. A.
The Astromaterials X-Ray Computed Tomography Laboratory at Johnson Space Center **[#6061]**
The Astromaterials Curation Office has recently begun incorporating X-ray CT data into the curation processes for lunar and meteorite samples, and long-term curation of that data and serving it to the public represent significant technical challenges.

1:50 p.m. Wagner R. V. * Robinson M. S.
PitScan: Computer-Assisted Feature Detection **[#6051]**
We developed PitScan to assist in searching the very large LROC image dataset for pits — unusual <200m wide vertical-walled holes in the Moon's surface. PitScan reduces analysts' workload by pre-filtering images to identify possible pits.

2:10 p.m. Mahanti P. * Lanjewar K. Thompson T. LROC Team
Classification of Small Lunar Crater Morphological State by Deep Learning **[#6028]**
Deep learning methods are used to classify small lunar impact craters by their degradation states from LROC images.

2:30 p.m. Parente M. * Gemp I.
Raman/LIBS Data Fusion via Two-Way Variational Autoencoders **[#6064]**
We propose an original solution to extracting mineral abundances from Raman spectra by combining Raman data with LIBS using a novel deep learning model based on variational autoencoders and data fusion, which outperforms the current state of the art.

2:50 p.m. Padams J. * Grimes K. Hollins G. Lavoie S. Stanboli A. Wagstaff K.
NASA PDS IMG: Accessing Your Planetary Image Data **[#6034]**
The Planetary Data System Cartography and Imaging Sciences Node provides a number of tools and services to integrate the 700+ TB of image data so information can be correlated across missions, instruments, and data sets and easily accessed by the science community.

3:10 p.m. BREAK

3:40 p.m. Paganelli F. * Conrad A.
Ground Observation of Asteroids at Mission ETA **[#6032]**
We focused on Lucy's targeted asteroids to derive information for best ground-based observation at mission ETA. We used a workflow for data extraction through JPL Horizons considering the LBT-MODS 1. Results outline opportunities suitable during close approach of Lucy ETA.

4:00 p.m. He L. * Arvidson R. E. O'Sullivan J. A.
Retrieving Single Scattering Albedos and Temperatures from CRISM Hyperspectral Data Using Neural Networks **[#6052]**
We use a neural network (NN) approach to simultaneously retrieve surface single scattering albedos and temperature maps for CRISM data from 1.40 to 3.85 μm. It approximates the inverse of DISORT which simulates solar and emission radiative streams.

4:20 p.m. Oyen D. A. * Komurlu C. Lanza N. L.
Interactive Gaussian Graphical Models for Discovering Depth Trends in ChemCam Data **[#6060]**
Interactive Gaussian graphical models discover surface compositional features on rocks in ChemCam targets. Our approach visualizes shot-to-shot relationships among LIBS observations, and identifies the wavelengths involved in the trend.

4:40 p.m. Raimalwala K. Faragalli M. Reid E.
Autonomous Soil Assessment System: A Data-Driven Approach to Planetary Mobility Hazard Detection **[#6036]**
The Autonomous Soil Assessment System predicts mobility hazards for rovers. Its development and performance are presented, with focus on its data-driven models, machine learning algorithms, and real-time sensor data fusion for predictive analytics.

CONTENTS

THE PYTHON SPECTRAL ANALYSIS TOOL (PYSAT) FOR POWERFUL, FLEXIBLE, AND EASY PREPROCESSING AND MACHINE LEARNING WITH POINT SPECTRAL DATA R.B. Anderson[1], N. Finch[1], S. Clegg[2], T. Graff[3], R.V. Morris[3], J. Laura[1]; [1]U.S. Geological Survey, Astrogeology Science Center, Flagstaff, AZ (rbanderson@usgs.gov); [2]Los Alamos National Laboratory, [3]NASA Johnson Space Center.

Introduction: Many planetary spectroscopy instruments such as ChemCam, SuperCam, Alpha Particle X-Ray Spectrometer (APXS), Mossbauer, Planetary Instrument for X-ray Lithochemistry (PIXL), Mini-TES, etc. collect point spectral data. Interpretation of these data is vital to the understanding of the geology of the targets and sites analyzed, but point spectral data are considerably more difficult to work with than images. Even for members of instrument teams, it can be challenging to apply new processing and analysis techniques to the data.

We have developed the free and open-source Python Spectral Analysis Tool (PySAT) library and point spectra interface to enable the planetary community to process and analyze point spectra without requiring programming expertise. This work is distinct from, but complimentary to, PySAT development geared toward orbital imaging spectrometer data [1].

Data Format: PySAT uses the Pandas library [2] to efficiently store spectra and associated metadata in a single data frame. The primary format for point spectral data is a comma-separated value (.csv) file with spectra and associated metadata stored in rows. The .csv file has two-level column labels with the top row indicating broad categories of data (e.g. wavelength, metadata, composition) and the second row indicating specific categories (e.g. '240.811', 'Target Name', 'SiO_2'). The PySAT point spectra tool can read individual PDS-format "clean, calibrated spectra" (CCS) files from ChemCam into the PySAT .csv format. We also provide laboratory data from a Laser-Induced Breakdown Spectroscopy (LIBS) instrument at Johnson Space Center in the proper format to be used with PySAT.

Capabilities:

Preprocessing: The PySAT point spectra tool provides a number of useful preprocessing capabilities. Once loaded, data frames can be manipulated by removing rows, splitting a single data frame, merging multiple data frames into one, multiplying all spectra in a data frame by a vector, or finding the derivative of each spectrum. Spectra in one data frame can also be interpolated onto the spectral channels from another data frame, a first step in combining data from different instruments.

A mask, specified by a simple .csv file, can be applied to all of the spectra in a data frame, and spectra can be normalized to the integrated signal within specified wavelength range(s). Spectra can also be grouped into a number of "folds" in preparation for model validation and testing, with the folds stratified on a single column of the data frame's metadata (for example, the SiO_2 content) to ensure a similar distribution of that variable in each fold. PySAT leverages the scikit-learn machine learning library [3] to enable several different dimensionality reduction methods: principal components analysis (PCA), two different independent component analysis (ICA) algorithms, as well as t-distributed stochastic neighbor embedding (t-SNE), and Locally Linear Embedding (LLE). The PySAT library also includes a number of continuum removal algorithms provided by [4]. We have also implemented several outlier removal methods, including Isolation Forest, Local Outlier Factor, One-class SVM, and Elliptic Envelope.

We have also added a "peak area" method which identifies local minima and maxima in the spectra, sums the signal between the minima, and saves the new value under the wavelength of the maximum. This effectively collects the signal from a peak into a single channel, shrinking the data set. This has been shown to improve calibration based on weak emission peaks [5]. Future updates will include true peak-fitting capabilities.

Regression: Regression methods can be used to generate models capable of converting observed spectra into predictions of target properties, such as chemical composition. PySAT leverages scikit-learn [3] to allow users to train a variety of regression models, including: Automatic Relevance Determination (ARD), Bayesian Ridge Regression (BRR), Elastic Net, Gaussian Process Regression, Least Angle Regression (LARS), least absolute shrinkage and selection operator (Lasso), Ordinary Least Squares (OLS), Orthogonal Matching Pursuit (OMP), Partial Least Squares (PLS), Ridge regression, and Support Vector Regression (SVR). To assist users in identifying the optimal parameters for these methods, a flexible cross validation option is available using the stratified folds defined in preprocessing.

Once a regression model has been trained, it can be used for prediction. If multiple models have been trained, they can be combined to implement submodel regression [6]. The current ChemCam calibration uses blended PLS submodels [7], but there is no requirement in PySAT that the same regression method be used for all submodels, providing added flexibility. The ranges over which the submodels are blended in

PySAT can optionally be optimized based on performance on training data.

Visualization: PySAT also includes Matplotlib-based [8] options for producing point and line plots, as well as plots of scores and loadings to visualize PCA and ICA results. Figure 1 shows some example plots.

Interface: We have also developed a graphical user interface based on PyQt5 [9] to make all of the above capabilities accessible for non-programmers. The graphical interface is based on the concept of "workflows", comprising individual processing steps or "modules" arranged in a user-specified order. The interface includes a progress bar to indicate when the program is running a calculation. As each module is run it is grayed-out, indicating progress through the workflow. A console view at the bottom of the interface prints messages as the modules run. The interface allows users to re-run or delete the last module, and stop the run partway through.

Once the user has determined the optimal order for the steps in the workflow, it can be saved and restored at a later date. This saving and restoration process uses Python's "Pickle" capability, and provides a convenient way to store and later replicate the exact processing steps involved in deriving results from a data set. For example, the methods and plotting commands used to produce the results and figures in a publication could be stored and included as supplemental information along with the manuscript.

Future Work: Both the PySAT point spectra tool GUI [10] and the underlying code containing the core functionality of PySAT for point spectra [11] are available on github. As development continues, several additional capabilities will be added, including: clustering and classification; calibration transfer; and interactive plotting. The course of development and prioritization of different features will also be guided by feedback from users. We will also work to fully document the tool and develop tutorials.

Although the tool was designed primarily based on analysis of LIBS spectra, it can be easily applied to any other spectral data. The tool provides a flexible and powerful framework enabling scientists to process and analyze spectral data using multivariate and machine learning methods, leading to improved scientific results.

References: [1] Gaddis et al. (2017) Planetary Data Workshop, #7060 [2] http://pandas.pydata.org/ [3] http://scikit-learn.org/ [4] Giguere, S., Carey, C.J., Boucher, T., et al. (2013) Proc. 5th IJCAI Workshop on Artificial Intelligence in Space. [5] Clegg, S.M. et al. (2017) LPSC XLVIII, #2439 [6] Anderson, R.B., et al. (2017) Spectrochim. Acta B, 129, 49–57. doi:http://dx.doi.org/10.1016/j.sab.2016.12.003 [7] Clegg, S. et al. (2017), Spectrochim. Acta B, 129, 64-85. http://doi.org/10.1016/j.sab.2016.12.003 [8] http://matplotlib.org/ [9] https://riverbankcomputing.com/software/pyqt/intro [10] https://github.com/USGS-Astrogeology/PySAT_Point_Spectra_GUI [11] https://github.com/USGS-Astrogeology/PySAT

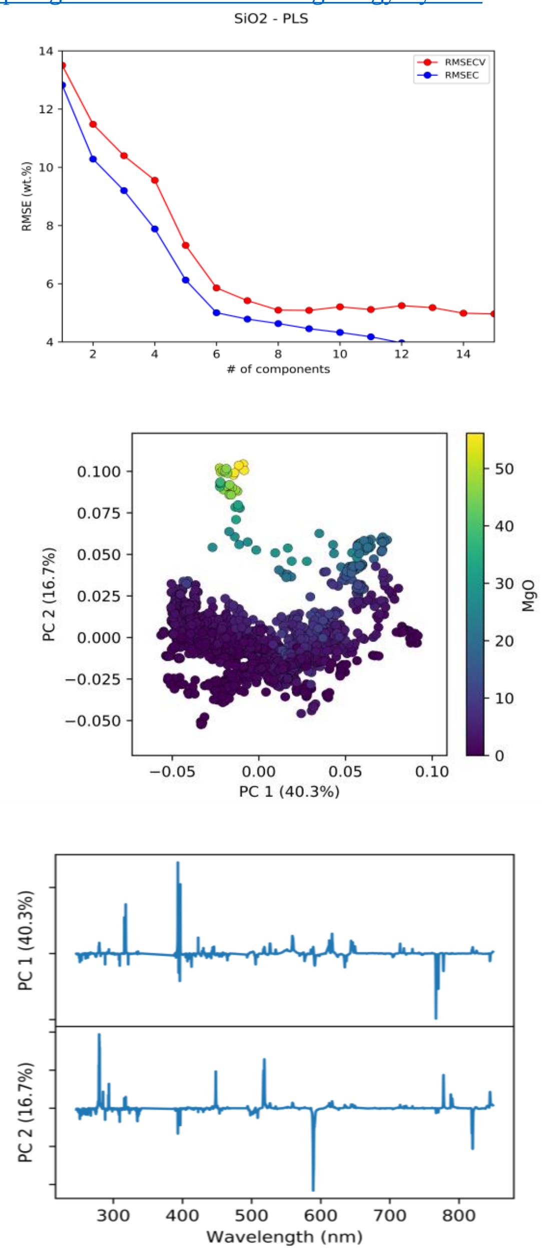


Figure 1: *Top*: Cross validation results for a PLS model of SiO2 based on LIBS spectra. *Middle*: Plot of the PCA scores for the first two principal components of a database of LIBS spectra. Points are color coded according to MgO content. *Bottom*: The corresponding PCA loadings for the first two components.

PLANETARY COORDINATES RECOMMENDATIONS FROM THE IAU WORKING GROUP ON CARTOGRAPHIC COORDINATES AND ROTATIONAL ELEMENTS. Brent Archinal[1] and the IAU Working Group on Cartographic Coordinates and Rotational Elements. [1]U. S. Geological Survey, 2255 N. Gemini Drive, Flagstaff, AZ 86001, USA barchinal@usgs.gov.

Overview: Community standards are an important component of Planetary Spatial Data Infrastructure [1] and therefore of planetary informatics and data analytics. Since 1979, the Working Group on Cartographic Coordinates and Rotational Elements (hereafter the "WG") of the International Astronomical Union (IAU) has, after most IAU General Assembly (GA) meetings, issued a report recommending coordinate systems and related parameters (body orientation and shape) that can be used for making cartographic products (maps) of solar system bodies. These recommendations, which are open to further modification when indicated by community consensus, are intended to facilitate the use and comparison of multiple datasets by promoting the use of a standardized set of mapping parameters. This abstract is intended to draw attention to the WG's efforts, our previous reports, and our just published (following the 2015 GA) report [2]. This effort provides support for the many fields covered by this conference, such as interoperability, data modeling, data visualization and interpretation, and planetary data processing generally. The WG encourages input and is available to assist users, instrument teams, and missions. See our website [3] for additional information.

Operation of WG: The Working Group consists of 18 volunteers, including C. Acton, B. Archinal (Chair), A. Conrad (Acting Vice Chair), G. Consolmagno, T. Duxbury, D. Hestroffer, J. Hilton, L. Jorda, R. Kirk, S. Klioner, D. McCarthy, K. Meech, J. Oberst, J. Ping, K. Seidelmann, D. Tholen, P. Thomas, and I. Williams. Our most recent report included substantial input from the late M. A'Hearn. Volunteers may join the WG at any time, and usually join for at least a three-year term to help with each new report following the IAU GA. The WG looks at new determinations of coordinate systems (e.g., body sizes and orientations) that preferably have been published in refereed papers, and makes recommendations as to which to use, based where possible on consensus decisions. As a volunteer organization, the WG has no resources to verify results or conduct its own research so it relies only on published results and community input. For that reason, it is sometimes not possible to recommend one set of results over another. The WG cannot verify or "bless" any particular results. The WG has no "enforcement" powers, but tries, in reflecting the long-term planetary community consensus, to make persuasive recommendations. The WG does not deal with issues related to the formats of mapping products. Such issues have largely been left to individual map developers, archiving organizations such as the NASA Planetary Data System (PDS), the IPDA (International Planetary Data Alliance), or the NASA Mars Geodesy and Cartography and Lunar Geodesy and Cartography Working Groups (MGCWG [4], LGCWG [5]), individual missions, and the NASA Mapping And Planetary Spatial Infrastructure Team (MAPSIT) [6]. Input from such organizations has been welcomed by the WG and the frequency of interaction highlights the strong need for such organizations at mission, space agency, and international levels.

As pointed out at the 2012 IAU GA [7] a substantial body of IAU recommendations exist that have been developed over many decades of input by IAU members, national space agencies, and other institutions. Care should be taken to follow such recommendations or to present well-reasoned arguments why they should be changed. The IAU and its Working Groups stand ready to help authors, journal editors, and missions understand and follow IAU recommendations.

Defining Longitude: One recent issue is the question of how the definition of longitude should be updated on Solar System bodies. The WG addressed this issue in its first report [8] and reiterates in our new report [2] that once an observable reference feature at a defined longitude is chosen, the longitude definition origin should not change except under unusual circumstances (such as perhaps a change in or loss of the feature). Given that our definition of longitude is primarily for mapping surface features, it is more logically tied to data related to the surface of the body (e.g., direct imaging or altimetry) than to dynamical data (e.g., the principal axes of inertia for resonantly or synchronously rotating bodies such as Mercury [9], the Moon, or Jovian or Saturnian satellites). Once such a feature has been adopted, changing to a longitude system defined by some other method should be avoided. Note that this recommendation does not preclude the use of smaller or more precisely determined features, multiple features, or even human artifacts to define longitude, as long as the original definition is maintained to the level of precision at which the feature can be located in new data. Some shift in longitude of previously identified features may occur whenever new data are available and processed, but this is minimized at least in the vicinity of the defining feature.

Coordinate System for (4) Vesta: In August, 2011, the NASA/DLR/ASI Dawn mission proposed using a longitude system with a large (~155°) rotation from the previous [10] system. Many reasons were expressed for this new system, but the WG replied in both September 2011 and March 2012, after careful and extensive consideration, that the arguments were not compelling enough to ignore previous usage by the planetary community and the WG's previous recommendations. Unfortunately, the mission began publishing results using only their rotated system. This resulted in substantial confusion. Fortunately, the NASA Planetary Data System requires that data products it archives follow various international and NASA standards, including those of the IAU. The mission therefore proposed a new system, which the PDS did accept as agreeing with IAU recommendations. This system is as described in the archive [11] (with W_0=285.39º). The WG ultimately formally accepted this system and recommended it for general use [2, 12].

General Changes: Following extensive discussion, substantial updates have been incorporated by the WG into our new report. An overview follows. *First*, based on the experience with Vesta, the WG has reworded and clarified its recommendations regarding updating longitude. *Second*, mission and community input indicates a need for the WG to differentiate between planetary body shapes and sizes for image projection and scientific modeling vs. a reference surface for elevation and map scale. In particular, long-accepted values for the latter are documented for the Moon and (now recommended for) Titan. *Third*, after considerable input from the community, including from New Horizons mission personnel, the discussion of terminology for the poles (hemispheres) of small bodies has been modified, e.g. to indicate that following community practice, cardinal directions can still be used informally or as shorthand for directions on small bodies (which formally have only positive and negative directions). *Fourth*, updates to the orientation models of Jupiter and Saturn are not recommended at this time, as we await community consensus on a model for Jupiter and final results from the Cassini mission regarding the orientation of Saturn.

Changes for Specific Bodies Under Discussion: Formulas for the Earth's orientation (which were previously given for comparison purposes only) have been removed in order to avoid confusion over their accuracy. The MGCWG has recommended a new orientation model for Mars (T. Duxbury, memo of 2017 August 18), which the WG in turn has recommended for use. More precision in longitude is provided by fixing the position of the Viking 1 lander. Neptune's rotation model has been updated based on new results from Karkoschka [13]. Individual members of the WG worked with Dawn mission personnel to arrive at a suitable way to update the existing orientation model for Ceres. New or updated orientation values are recommended for (52) Europa, (511) Davida, and (2867) Šteins. The declination of the pole of (243) Ida has been corrected. Orientation data were added for comet 9P/Tempel 1 based on the Stardust NExT flyby [14], for 19P/Borrelly based on the DS1 flyby and subsequent ground-based measurements [15], for 103P/Hartley 2 based on the EPOXI flyby [16], and for 67P/Churyumov-Gerasimenko based on the pre-perihelion approach mapping from the Rosetta orbiter [17]. Data for Mercury, (1) Ceres, and the radii for (134340) Pluto and Charon [18] have also been updated based on recent mission results and papers. The size of the Sun was updated per a 2015 IAU Resolution and sizes are given for (16) Psyche and (52) Europa, and the size of (25143) Itokawa has been corrected.

Other recommendations: We repeat our previous recommendations that planning and efforts be made to make controlled cartographic products; and newly recommend that common formulations should be used for orientation and size and that historical summaries of the coordinate systems for given bodies should be developed. We point out that for planets and satellites planetographic systems have generally been historically preferred over planetocentric systems, and that in cases when planetographic coordinates have been widely used in the past, there is no obvious advantage to switching to the use of planetocentric coordinates.

Request for Input: The WG desires continued input from the planetary community, especially regarding the systems for specific bodies, the operation of the WG, our proposed question submitting process, and posting of updates via our website. We also welcome volunteers to become WG members and help with our efforts. Our membership is open to all. The lead author of this abstract should be considered the primary point of contact.

Acknowledgements: Funding for B. Archinal has been provided via a NASA-USGS Interagency Agreement on Planetary Spatial Data Infrastructure.

References: [1] Laura et al. (2017) *ISPRS Int. J. Geo-Inf*, 6, 181. [2] Archinal et al. (2017) *CMDA*, in press. [3] http://astrogeology.usgs.gov/groups/IAU-WGCCRE. [4] Duxbury et al. (2002) *ISPRS, 34*, pt. 4, http://astrogeology.usgs.gov/groups/ISPRS. [5] Archinal and the LGCWG (2009) *LPS XL*, Abstract #2095. [6] Radebaugh et al. (2017), LEAG Annual Meeting, Abstract #5053. [7] Meech et al. (2012) *Inquires of Heaven*, no. 10, p. 6, [8] Davies et al. (1980) *Celest. Mech., 22*, 205-230. [9] Margot (2009) *CMDA, 105*, 329-336, DOI: 10.1007/s10569-009-9234-1. [10] Thomas et al. (1997) *Icarus* 128, 88–94; also see [2]. [11] DAWN mission, (2012) http://sbn.psi.edu/archive/dawn/fc/DWNVFC2_1A/DOCUMENT/VESTA_COORDINATES/VESTA_COORDINATES_120918.PDF. [12] Archinal et al. (2013) http://astropedia.astrogeology.usgs.gov/download/Docs/WGCCRE/IAU-WGCCRE-Coordinate-System-for-Vesta.pdf. [13] Karkoschka, E. (2011) *Icarus*, 215, 439. [14] Belton et al. (2011) *Icarus* 213, 345. [15] Soderblom et al. (2004) *Icarus, 167*, 4; Mueller et al. (2010) *Icarus, 209*, 745. [16] Thomas et al. (2013) *Icarus 222*, 550; Belton et al. (2013) *Icarus, 222*, 595; Farnham and Thomas (2013) DIF-C-HRIV/MRI-5-HARTLEY2-SHAPE-V1.0, NASA PDS. [17] Preusker et al. (2015) *A&A*, 583, A33; Scholten et al. (2015) RO-C-MULTI-5-67P-SHAPE-V1.0:CHEOPS_REF_FRAME_V2, NASA PDS and ESA PSA. [18] Nimmo et al. (2017) *Icarus* 287, 12.

ESA PLANETARY SCIENCE ARCHIVE ARCHITECTURE AND DATA MANAGEMENT.

C. Arviset[1], I. Barbarisi[2], S. Besse[3], M. Barthelemy[4], G. de Marchi[5], R. Docasal[3], D. Fraga[3], E. Grotheer[2], D.Heather[1], C. Laantee[1], T. Lim[4], A. Macfarlane[2], S. Martinez[1], A. Montero[2], J. Osinde[4], C. Rios[2], J. Saiz[6], and C. Vallat[6], [1]ESA-ESAC, Camino Bajo del Castillo s/n, Urb. Villafranca del Castillo, 28692 Villanueva de la Canada, Madrid, Spain (Christophe.Arviset@esa.int, Caroliina.Laantee@esa.int), [2]SERCO for ESA-ESAC (Isa.Barbarisi@esa.int, Alan.Macfarlane@esa.int, Angel.Montero@esa.int, Carlos.Rios@esa.int), [3]AURORA for ESA-ESAC (Sebastien.Besse@esa.int, Ruben.Docasal@esa.int), [4]TPZ VEGA for ESA-ESAC, [5]ESA-ESTEC, Keplerlaan 1, Noordwijk, The Netherlands, (gdemarchi@cosmos.esa.int), [5]RHEA for ESA-ESAC (Jose.Osinde@esa.int, Jaime.Saiz@esa.int, Claire.Vallat@esa.int).

Introduction: The Planetary Science Archive (PSA) is the European Space Agency's (ESA) repository of science data from all planetary science and exploration missions. First released in 2004, it went through a completely renewed user interface and system architecture, with the "new" PSA v5.0 released early 2017 and can be accessed at http://psa.esa.int.

The PSA is being developed and operated within the ESAC Science Data Centre (ESDC) in collaboration with ESA's planetary Science Operations Centres, also located at the European Space Astronomy Centre near Madrid, Spain and with the missions' instrument teams mostly in Europe. The ESDC also hosts other science archives for ESA space science astronomy and heliophysics missions (http://archives.esac.esa.int/).

PSA content: The PSA is a multi mission archive containing science datasets from ESA's active planetary missions (Rosetta, Mars Express and Exomars TGO), legacy missions (Venus Express, Huygens, Giotto and Smart-1) and will include science data holdings from future missions (BepiColombo, Exomars RSP and Juice).

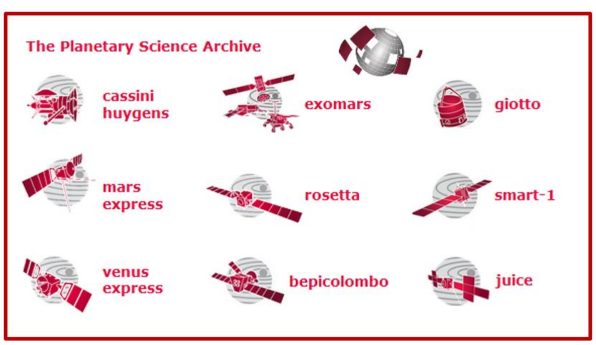


Figure 1: ESA Planetary Missions

All data sets are scientifically peer-reviewed by independent scientists, and are compliant with the Planetary Data System (PDS) standards. From PSA v5.0 early 2017, the archive can ingest datasets from PDS3 and PDS4 and offer them to scientists through common interfaces.

PSA interfaces: The primary way to access ESA planetary science data holdings is through the PSA GUI, which offers a powerful and user friendly faceted search web interface (Figure 2: PSA web GUI). Lots of work has been done to ensure homogenous metadata across the many instruments, to enable science driven searches across instruments across missions.

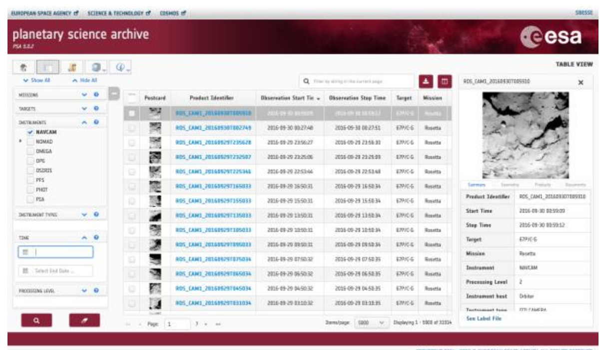


Figure 2: PSA web GUI

Most of the data is public and can be downloaded directly. Login is required for proprietary data download for authorized users, but most of the data is public and accessible without requiring to login.

Although planetary scientists represents the main users of the PSA, there are many spectacular images that can also be of interest of the general public and the media. To facilitate access to these, an archive image browser (Figure 3: PSA Archive Image Browser) is also available to quickly search and visualize Rosetta browse products. Plans for a map browser using GIS technologies are on-going, based on a still to be defined uniformed geometry information across missions datasets.

Users more experts with the ESA planetary missions can directly download full instruments datasets from the PSA FTP server.

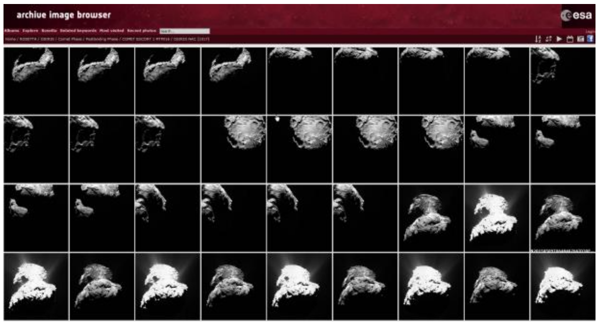


Figure 3: PSA Archive Image Browser

The new PSA also attempts to increase interoperability with the international community by implementing recognised planetary science specific protocols such as the PDAP (Planetary Data Access Protocol) and EPN-TAP (EuroPlanet-Table Access Protocol).

PSA Architecture: The PSA is based on a modular and flexible 3-tier architecture (Figure 4: PSA System Architecture). The storage layer consists of a data repository (around 45TB) and a RDBMS (PostgreSQL). The Client layer offers the main interfaces to scientists, either through a thin web client or through command line interface. The Server layer takes care of handling all database queries, data distribution services and all archive administrative services (ie user login and authentication, usage statistics).

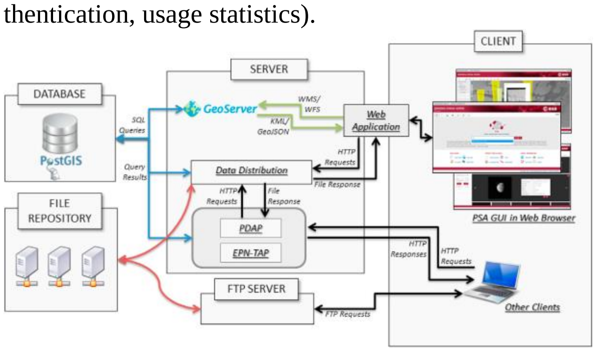


Figure 4: PSA System Architecture

PSA and IPDA: The International Planetary Data Alliance (IPDA) is an association of partners worldwide with the aim of improving the quality of planetary science data and services to the planetary science community (http://www.planetarydata.org). IPDA's mission is to facilitate global access to, and exchange of, high quality scientific data products managed across international boundaries. Ensuring proper capture, accessibility and availability of the data is the task of the individual member space agencies. Through the PSA team, ESA has been a co-founder of the IPDA with NASA and PDS and continues to be an active members of the alliance.

PSA's interoperability with other planetary data archives is obtained through the usage of common data formats (PDS3 and PDS4) and the implemention of common data interoperability protocols (ie PDAP and more recently EPN-TAP).

References:

[1] Besse, S. et al. (2017) *Planetary and Space Science*, doi, ESA's Planetary Science Archive: Preserve and present reliable scientific data sets.

[2] Macfarlane A. et al. (2017) *Planetary and Space Science*, doi, Improving accessibility and discovery of ESA planetary data through the new Planetary Science Archive.

Acknowledgments: The PSA team is very grateful to all the people who have contributed in some way for the last decade to the previous and current versions of the ESA's Planetary Science Archive. We are also thankful to the ESA's teams who are operating the missions (Science Operations Centres and Mission Operations Centres) and to the instrument science teams who are generating and delivering scientific calibrated products to the archive.

SUCCESSFUL DESIGN PATTERNS IN THE DAY-TO-DAY WORK WITH PLANETARY MISSION DATA.
K.-M. Aye, Laboratory for Atmosphere and Space Physics, University of Colorado Boulder, 1234 Innovation Drive, Boulder, CO 80303 (michael.aye@lasp.colorado.edu)

Introduction: After several years of processing and analyzing data from solar system missions and Citizen Science data I have recognized a few repeating patterns in the supporting software I created for working with these data.

Specifically, for the case of explorative data analysis, where it is often necessary to study and compare data from different processing or quality levels, or between the new derived product and the incoming data, three support structures turned out to be the most helpful patterns. These most prevalent patterns are what I call path managers data (or database) managers, and metadata searches.

For data processing, I have established a few simple tools that enable me to execute embarrassingly parallel workloads in parallel on multi-core machines without getting into parallel programming details.

Data(base)Manager: Often one has to work with different sources of data for the same analysis, either because one production set is newer and/or not published at the same source, or because one set has been produced as a derivative from the other. Other use cases include managing different versions of data dumps where for regression checking it is important that the same code can be executed easily with different sets of incoming data.

In these cases, it is very helpful to define a software tool (a Python class in my case) that helps with easy finding and access to the different data.

The characterizing features are:

1) distinguish between and provide different versions of datasets in the project storage folder; 2) node/machine-dependent storage locations (e.g. laptops usually have less capacity than desktops); 3), sub-dataset retrieval functions that are able to filter the currently active data-set for keywords and/or ranges of parameters, easily providing a subset for the ongoing work task.

Node/machine-dependent storage locations can be managed by a configuration file, that, once created, are used by the DataBaseManager for looking up the relevant paths for the current machine. This provides the user with abstraction from the local storage, while the interactive data-analysis focuses on receiving a data-object of interest. An example use looks like this:

db = DBManager() # using paths from config file

or

db = DBManager(temp_path) # using temporary paths to test a new dataset.

Finally, to receive data of interest, I simply do:

data = db.get_file_id(id)

This design abstracts me from distracting and disturbing long path management.

PathManager: Hand-in-Hand with the DataBaseManager, I employ what I call PathManagers. While sounding similar to manage database storage paths (those are managed by config files), this class is designed to help with data products during complex and long processing pipelines. Each project will have, either defined by others, or by the researcher themselves, a structure of paths where different levels of data, or metadata data for the actual data are stored.

For spacecraft missions this results in a large list of possible paths below the database path, for each and every data sub-product and it is very time-consuming to manually keep track of these while performing real-time explorative data analysis.

In recent projects I have implemented PathManagers to deal with this problem.

Basically, a (currently hard-coded) class structure is defined that has attributes for each kind of sub-data paths underneath a common observation id. Often, the actual file paths are named in strategic ways to enable alpha-numeric sortability, but those are usually not supportive for real-time data analysis, for being hard to remember or overly structured. The PathManager enables an abstraction layer between how products are stored on disk (=more structured) and how they are efficiently accessed during interactive data analysis (=more memorizable).

For example, for a specific HiRISE observation ID, there are paths to many products related to one observation ID, like COLOR, RED, and combined mosaics.

During explorative analysis, I am able to call up a PathManager object for a given OBSID and database, and an PathManager attribute called "red_mosaic" would provide the full path to that image product simply like so:

pm = PathManager(obsid, db) # receiving a DataBaseManager from above

print(pm.red_mosaic)

pm.red_mosaic now contains a long path starting from basic storage as managed by the DataBaseManager object, but also, inside that, for the given OBSID, a complex file name construct indicating the access path to the MOSAIC made from only the RED products, inside that database, under the given

OBSID folder. The beauty of it, again, is that all the distracting path management has been hidden and the user can focus on providing the path object to the next tool in the processing chain.

Metadata searches: Finding data often involves the search through metadata. To enable this without using often cumbersome and mostly un-controllable web interfaces, the Planetary Data System archive requires the delivery of a cumulative index file that summarizes a chosen set of metadata for each observation id of a data set.

I have found a way to ingest these metadata tables into a pandas DataFrame for convenient search queries. However, one current problem is that each delivery to any of the PDS nodes comes with a new cumulative index file that is stored into a new subfolder.

This creates problems for tool creators like me that want to provide the newest metadata for a given mission instrument.

It basically would create the need to parse the html code for the most recent folder that can be found which could be done at a hacking session at this conference, but recently, the Rings-Moons node has implemented my suggesting to create these static URLs that will link to the most recent delivered cumulative index file.

Using the remote database API that is provided by the PDS Rings-Moon node I created a tool that enables me to plug in an IMAGE_ID from a paper I read, retrieve the image, store it automatically at the right place using the above described DatabaseManager, and start a recalibration processing chain, if required.

Abstraction: For my efforts in creating a tool-set for planetary science in Python, similar to what \`astropy\` did for astrophysics, I believe it would be helpful to abstract these patterns into either a) configurable templates that should be easy adaptable for any new projects, so that it could become either part of the \`planetpy\` package, or b), become part of a so called “cookiecutter” project template that could be used to create a new software package for a new science project.

I will provide concrete suggestions on how this could be done and am asking for community feedback on the best way forward.

The above mentioned metadata approach using pandas \`DataFrames\` is ready to be implemented into the \`planetpy\` project, it might just be more or less cumbersome to add the most recent cumulative index file to the current project, depending on the willingness of the PDS node to provide static links to these files.

Parallel processing: An often occurring case is the application of the same function on a large set of images or other data products. As these operations all occur in the subfolder of a given data product, these operations are all independent and can be executed in a parallel fashion without creating race conditions. These workloads are defined as being “embarrassingly parallel” and it is still not the case that available analysis suites provide easy-to-use toolsets to process these automatically across all cores on a multi-core desktop PC. Using the parallelization modules of IPython [1], I have implemented a few small tools that enable me to execute a function on a list of products easily, without having to think about setting up or starting parallel engines and such. This tool also provides feedback via a graphical progress widget, as provided by the Jupyter notebook tools.

For the conference I will present and discuss these tools [2] using examples of my work with medium-sized Citizen Science databases and planetary imaging data.

References:

[1] Fernando Pérez, Brian E. Granger, IPython: A System for Interactive Scientific Computing, Computing in Science and Engineering, vol. 9, no. 3, pp. 21-29, May/June 2007, doi:10.1109/MCSE.2007.53. URL: http://ipython.org

[2] K.-Michael Aye, & jocu8995. (2016, November 11). michaelaye/pyciss: PathManager and databases (Version v0.6.0). Zenodo. http://doi.org/10.5281/zenodo.166116.

THE SMALL BODY MAPPING TOOL (SBMT) FOR ACCESSING, VISUALIZING, AND ANALYZING SPACECRAFT DATA IN THREE DIMENSIONS. O. S. Barnouin[1], C. M. Ernst[1], R. T. Daly[1], and the Small Body Mapping Tool Team[1]. [1]The Johns Hopkins University Applied Physics Laboratory, 11101 Johns Hopkins Road, Laurel, MD, 20723, USA (sbmt@jhuapl.edu).

Introduction: Spacecraft missions return massive amounts of valuable data, but those data can be hard to access, visualize, and analyze. Most asteroids, comets, Kuiper belt objects, and small moons present additional challenges because two-dimensional map projections severely distort features on irregularly shaped bodies. The Small Body Mapping Tool (SBMT) developed at the Johns Hopkins University Applied Physics Laboratory addresses these challenges [1].

The SBMT lets users search for spacecraft data and project it onto shape models of small bodies. As a result, users can quickly find the data they need, look at the data in context, and do their science in three dimensions, without worrying about map projection issues or wading through the Planetary Data System (PDS) archives. Alternatively, the SBMT can be a starting point: users can pinpoint the data they need using the SBMT and then download the raw data from the PDS. The Tool includes a diverse suite of bodies and data types (images, spectra, altimetry data, see "Available Data") and supports co-registration of these data products. It has been or is being used by multiple mission teams, including Dawn, Rosetta, OSIRIS-REx, and Hayabusa2.

The Small Body Mapping Tool is publically available as a free download at sbmt.jhuapl.edu. It works on Mac, Linux, and Windows operating systems and has an easy-to-use graphical user interface. The SBMT is written in Java and uses the Visualization Toolkit (VTK), an open-source, freely available software system for 3D computer graphics, rendering, and visualization [2]. Some datasets and functionality, such as those for active missions, have restricted access; however, such features become publically available in the SBMT once the data have been archived with the PDS.

Features: The SBMT facilitates interactive searches for spacecraft data. This capability allows users to quickly and easily identify the images, spectra, or altimetry data that will help them achieve their science or engineering objectives. Once selected, data can be projected onto the shape model and analyzed using the SBMT's built-in analysis tools, thereby integrating data discovery and data analysis. Alternatively, users can export data for use in analysis tools of their choice.

The Tool's graphical user interface includes several tabs next to a large viewing area. Once users choose a body from a menu, each tab provides access to a different dataset. Users can set shape model illumination and simulate camera pointing. In the viewing area, users can interactively manipulate the shape (rotate, zoom, etc.).

Body tab: The body tab allows users to visualize a shape model at a variety of resolutions, view a basemap (where available), and overlay color maps of elevation, slopes, gravitational potential, and gravitational acceleration onto the shape. Such geophysical maps have proven useful in studies of asteroids [3–5].

Data tabs: Once users choose a body, the SBMT interface populates tabs based on the available data. Once a particular data tab is selected, users can search based on many parameters, including emission, incidence, and phase angles; pixel scale; data acquisition time; and wavelength. Users can also search for data by location by selecting a region of interest on the shape. The SBMT displays the footprints of images, spectra, and altimetry data found by the search so that users can decide which to load (Fig. 1). Users can simulate lighting to match the conditions when the data were acquired.

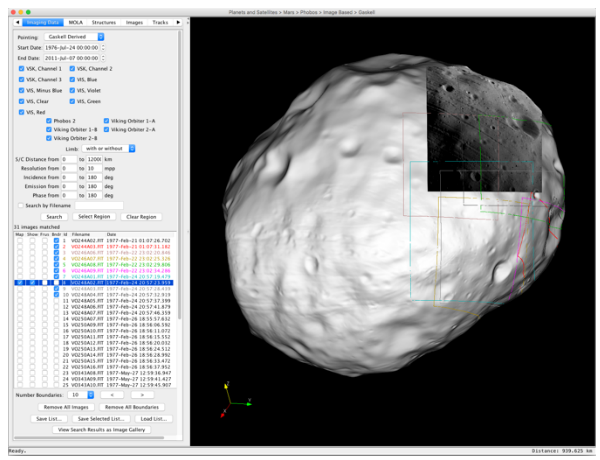


Figure 1. The SBMT lets users easily search for data. Here, available data footprints are shown (colored squares), and one image has been projected directly onto the shape model of Phobos.

The data tabs include several tools for data analysis. The functionality depends on the data type (e.g., capabilities for images differ from those for lidar tracks). For images, the SBMT can generate image cubes for overlapping images. These cubes can then be used to make RGB composites or be exported (e.g., to ENVI). The custom bands feature for spectral data allows users to do band math [e.g., 6]. For laser altimeters, transects can be used to measure topography (Fig. 2).

Structures tab: The structures tab lets users map features on the shape model even while viewing images or other data. Paths and polygons can be used to map lineaments, regions, and geologic units [e.g., 7–10]. Craters and blocks can be mapped with circles or ellipses [e.g., 11, 12] (Fig. 3). Points can be used to mark the locations of features. The data are saved as human-

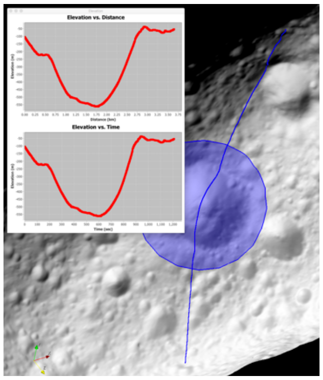


Figure 2. Topographic profiles can be extracted from lidar tracks and DTMs, such as this example from Eros.

readable XML (paths, polygons) or ASCII (circles, ellipses, points) files that can be easily imported into other tools or codes. The files contain both the measurements (e.g., polygon area, crater diameter) and regional geophysical data (e.g., elevation, slope).

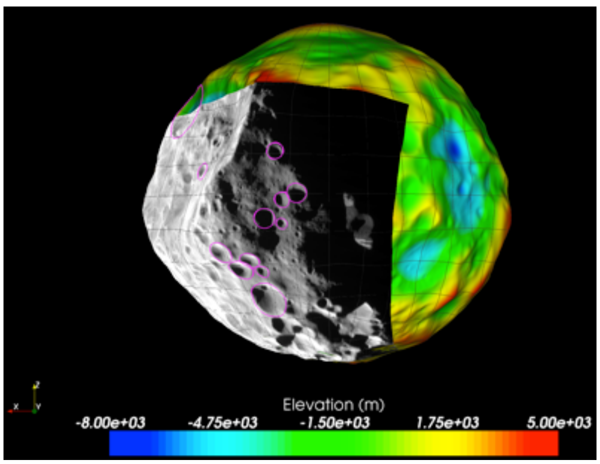


Figure 3. The SBMT lets users map craters, blocks, and other features directly on the shape. Here the user is mapping craters (magenta circles) on Phoebe. The shape is colored by elevation; a Cassini ISS image is draped on the shape.

Regional DTMs tab: For shape models generated using stereophotoclinometry [13], global shape models have lower resolution than the maplets on which the models are based. The regional DTMs (Digital Terrain Models) tab contains a database that allows users to construct higher-resolution regional DTMs, which can be overlain on the shape model or visualized independently in the SBMT. The SBMT allows users to collect topographic profiles across DTMs [e.g., 14].

Observing conditions tab: The observing conditions tab lets users visualize the relative positions of the spacecraft and the target body through time, including simulations of the lighting conditions and the sub-Earth, sub-spacecraft, and sub-solar points. Future enhancements to this tab will provide a way to link ground-based observations made at known times to specific parts of the object.

Custom data import: Each data tab allows users to import customized data and visualize it on the shape model. Users can apply pointing information from the SBMT to display imported data, as long as the files retain their original dimensions. Simple cylindrical global or regional basemaps can also be imported using this feature.

Available data: As of early 2018, the public version of the SBMT includes spacecraft data for several asteroids (Ceres, Vesta, Lutetia, Eros, Itokawa) and moons (Phobos, Dione, Mimas, Phoebe, Tethys). More bodies will soon be available, including an improved Phobos model, Deimos, 9P/Tempel 1, 67P/Churyumov-Gerasimenko, 81P/Wild 2, 103P/Hartley 2, and the saturnian moons Atlas, Calypso, Epimetheus, Helene, Hyperion, Janus, Pan, Pandora, Prometheus, Rhea and Telesto. The SBMT also works for large, spherical bodies like the Moon and Mercury [9, 15].

For each object, the SBMT includes a 3D shape model and available spacecraft datasets. Eros, for example, includes data from the NEAR-Shoemaker multispectral imager (MSI), near-infrared spectrometer (NIS), and NEAR laser rangefinder (NLR). Itokawa includes data from the LIDAR and AMICA imager.

In addition to those bodies detailed above, the SBMT has radar-, lightcurve, and image-based shape models for many objects, including those not yet visited by spacecraft.

Conclusion: The Small Body Mapping Tool is a powerful, easy-to-use tool for accessing and analyzing data from small bodies. The SBMT is actively being developed, and we will continue to release new datasets and functionality. Visit sbmt.jhuapl.edu to subscribe to the SBMT mailing list. We invite everyone in the community to reach out and discuss collaborations.

Acknowledgements: Much of the initial coding for the SBMT was done by E. Kahn at JHU/APL. In addition to the authors, many people have developed or supported the Tool, including D. Blewett, R. Klima, D. Buczkowski, J. Roberts, N. Chabot, L. Jozwiak, M. Zimmerman, J. Steele, L. Nguyen, J. Peachy, R. Turner, C. O'Shea, and A. Regeic. This work has been supported by grants from the NASA Planetary Science Division research and analysis programs.

References: [1] Kahn et al., 2011, *LPS 42*, abs. 1618. [2] Schroeder et al., 2006, *The Visualization Toolkit: An object-oriented approach to 3D graphics*, Kitware, Inc. [3] Cheng et al., 2002, *Icarus,* 155, 51–74. [4] Thomas et al., 2002, *Icarus,* 155, 18–37. [5] Barnouin-Jha et al., 2008, *Icarus,* 198, 108–124. [6] Klima et al., 2016, *LPS 47,* abs. 2572. [7] Buczkowski et al., 2008, *Icarus,* 193, 39–52. [8] Buczkowski et al., 2012, *GRL,* 39, L18205. [9] Ernst et al., 2015, *Icarus,* 250, 413–429. [10] Besse et al., 2014, *Planet. Space Sci.,* 101, 186–195. [11] Hirata, 2017, *Icarus,* 288, 69–77. [12] Mazrouei et al., 2014, *Icarus,* 229, 181–189. [13] Gaskell et al., 2008, *Met. Planet. Sci.,* 43, 1049–1061. [14] Roberts et al., 2014, *Met. Planet. Sci.,* 49, 1735–1748. [15] Deutsch et al., *Icarus,* 280, 158–171.

The Need for a Planetary Spatial Data Clearinghouse

Ross A. Beyer[1,2] Trent Hare[3] and Jani Radebaugh[4], [1]Carl Sagan Center at the SETI Institute, [2]NASA Ames Research Center, MS 245-3, Moffett Field, CA, USA (Ross.A.Beyer@nasa.gov), [3]Astrogeology Science Center, United States Geological Survey, and [4]Bringham Young University

As planetary missions become more diverse and mature, so does the data and the derived data created from those missions. This abstract focuses on planetary spatial data, which are inherently tied to locations on bodies in the solar system. It is information relative to specific places, and about those locations over time in a geographic and cartographic sense.

One of the aspects of a planetary spatial data infrastructure (PSDI) [1] is straightforward access to "science-ready" data. We lay out below how the concept of a planetary spatial data clearinghouse, similar to the National Geospatial Data Clearinghouse for Earth data, would put planetary sciences on a course to fully realize a PSDI.

The problem of finding Planetary Spatial Data: The Planetary Sciences community has a stable long-term archive in the form of the Planetary Data System (PDS). This archive is tasked with long-term storage, and serves as the cornerstone upon which a great deal of science and exploration have been accomplished over the decades. However, as the sophistication of science and exploration methods have increased over the years, users want more than just a static archive, they want more services than the PDS is currently able to provide.

Specifically in the realm of geospatial data, the planetary sciences community has a diversity of needs that it has tried to meet in various ways, by various actors, from various countries around the world. The wonderful thing is that they are not all trying to solve the exact same problem, but they are solving different problems or facets of problems that are all generally related to the concept of PSDI [1]. As a result, they are providing a vast array of valuable planetary spatial data.

The difficulty is when a researcher wishes to start a new endeavor, where do they start? Existing researchers have cobbled together knowledge of the complicated bazaar of planetary spatial data that exists, but even that knowledge is likely limited. New researchers or students could easily feel lost, and are in jeopardy of either spending time reproducing planetary spatial data that already exists or not using the best planetary spatial data they could, simply because they don't know that it exists.

Likewise, when space agencies begin large new projects, it is difficult for them to know what planetary spatial data is already available, and what they must fund to create.

How do Earth scientists deal with this issue? Earth scientists have very similar fundamental problems and issues regarding spatial data (Figure 1).

The Presidential Order that established the National Spatial Data Infrastructure of the United States [2] contained within it the concept of a National Geospatial Data Clearinghouse. And it is here that we may find a useful framework to think about in applying to our planetary problems.

The National Geospatial Data Clearinghouse does not, itself, primarily archive data. Instead, it acts as a central catalog of available data sets, through its `https://www.geoplatform.gov` portal, and those data sets are hosted elsewhere.

Providers of spatial data in earth sciences know that submission of their metadata to the National Geospatial Data Clearinghouse benefits their work (fulfilling the requirements of their funding bodies) and their colleagues. It makes their data searchable and findable, ultimately benefiting their field.

For more information about the National Geospatial Data Clearinghouse, the Federal Geographic Data Committee maintains an excellent concepts question and answer document at `https://www.fgdc.gov/dataandservices/clearinghouse_qanda`

What would this look like for planetary sciences? Laura et al. in [1] identify three foundational data themes for planetary sciences: geodetic coordinate systems, elevation, and orthoimages. Let's imagine a planetary spatial data clearinghouse which contains elevation and orthoimage data products for solar system objects. This clearinghouse would implement various database and web servers, but would be presented to a human user as a web site.

A researcher in need of spatial data comes to the clearinghouse looking for topography data and maps of Mars. When they search for topography, they would find the MOLA gridded data products, and may also find a merged HRSC/MOLA map. They might find regional terrain models created for landing sites, or even local terrain models created by researchers for their individual ROSES grants. When they search for orthoimages, they might find the USGS MDIM products, various USGS geologic maps, perhaps the excellent THEMIS mosaics hosted at ASU, or local mosaics and thematic maps that were created for published papers that are being hosted by the journal or the author's university.

The point here is that some of this data would be hosted in formal PDS archives, some might be hosted by large projects like ASU's JMARS data holdings [3], JPL's various Trek holdings [4], or even on data repositories at universities and institutions, but the clearinghouse doesn't hold or maintain the data, it simply acts as a "table of contents" to facilitate search and discovery which points back to where the actual data are hosted. The PDS provides an important archive of geospatial data, but there is far more "science ready" planetary spatial data available than exists in the PDS alone. A planetary spatial data clearinghouse would catalog and allow discovery of more than just the spatial data hosted by the PDS.

It is important to note that such a clearinghouse site must be managed to adapt and change. Its value would lie in the fact that there are human beings knowledgeable about solar system geospatial data that seek out published data in order to populate such a clearinghouse, and keep current on the field. Additionally, data providers could make their meta-data available via standards like the the OGC Catalog Services for the Web [5]. The growth and viability of such a clearinghouse requires both a managed search for data, and the ability for data providers to share the existence of their geospatial data.

A planetary spatial data clearinghouse as described above would be a valuable tool enabling planetary science research, and an important piece of realizing a planetary spatial data infrastructure.

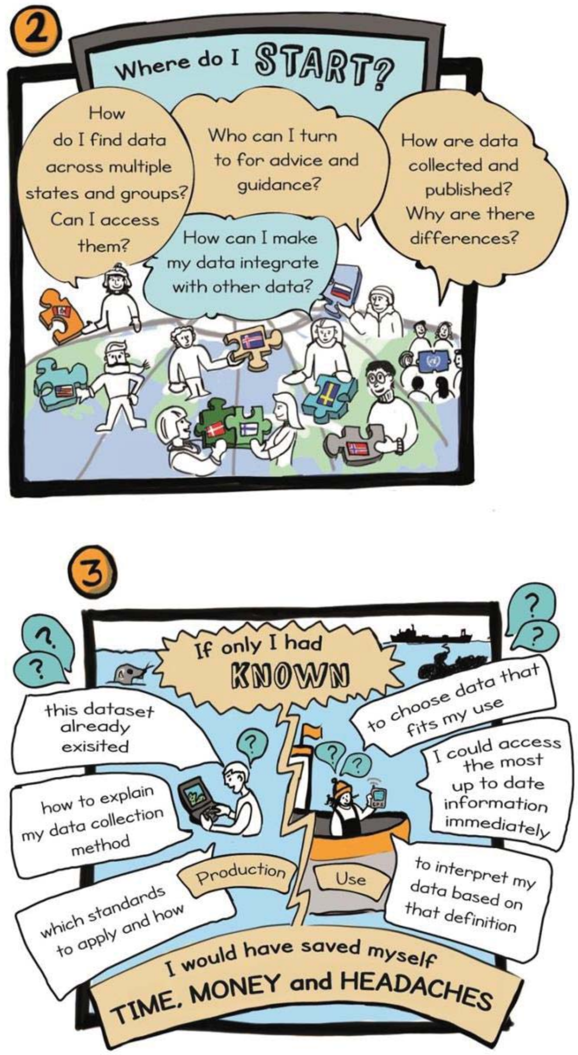


Figure 1: These cartoons are from the Arctic Spatial Data Infrastructure document [6] and are available under the Open Government License – Canada.

References: **[1]** J. R. Laura et al. "Towards a Planetary Spatial Data Infrastructure". In: *ISPRS International Journal of Geo-Information* 6.6 (2017), p. 181. DOI: 10.3390/ijgi6060181. **[2]** *Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure*. Exec. Order no. 12906, 59 Fed. Reg. 71 (April 11, 1994). 1994. **[3]** S. Dickenshied et al. "JMARS - Remote Sensing Visualization and Analysis for All Planetary Bodies". In: *LPI Contributions* 1986, 7126 (June 2017), p. 7126. **[4]** E. Law and B. Day. "Public Outreach with NASA Lunar and Planetary Mapping and Modeling". In: *European Planetary Science Congress* 11, EPSC2017-98 (Sept. 2017), EPSC2017–98. **[5]** T. M. Hare, L. R. Gaddis, and M. B. Bailen. "OGC Catalogue Services for Planetary Portals". In: *Lunar and Planetary Science Conference*. Vol. 46. Lunar and Planetary Inst. Technical Report. Mar. 2015, p. 2476. **[6]** Arctic SDI Board. *Spatial Data Infrastructure (SDI) Manual for the Arctic*. 1st ed. Sept. 2016. URL: https://arctic-sdi.org/wp-content/uploads/2017/04/SDI-Manual-for-the-Arctic-EDITED2_PS.pdf.

MASER: A Tool Box for Solar System Low Frequency Radio Astronomy. B. Cecconi[1], P. Le Sidaner[2], R. Savalle[2], X. Bonnin[1], P. Zarka[1], C. Louis[1], A. Coffre[3], L. Lamy[1], L. Denis[3], J.-M. Griessmeier[4], J. Faden[5], C. Piker[5], N. André[6], V. Génot[6], S. Erard[1], T. A. King[7], J. N. Mafi[7], M. Sharlow[7], J. Sky[8] and M. Demleitner[9].
[1] LESIA, Observatoire de Paris, PSL, CNRS, Meudon, France
[2] DIO, Observatoire de Paris, PSL, CNRS, Paris, France
[3] Station de Radioastronomie de Nançay, Obs. Paris, PSL, Univ. Orléans, CNRS, Nançay, France
[4] LPC2E, Univ. Orléans, CNRS, Orléans, France
[5] Univ. Iowa, Iowa-City, IA, USA
[6] IRAP, Univ. Paul Sabatier, CNRS, Toulouse, France
[7] IGPP, UCLA, Los Angeles, CA, USA
[8] Radio Sky Inc.
[9] Univ. Heidelberg, Heidelberg, Germany.

Introduction: The MASER (Measuring, Analysing and Simulating Radio Emissions) project provides a comprehensive infrastructure dedicated to low frequency radio emissions (typically < 50 to 100 MHz). The four main radio sources observed in this frequency are the Earth, the Sun, Jupiter and Saturn. They are observed either from ground (down to 10 MHz) or from space. Ground observatories are more sensitive than space observatories and capture high resolution data streams (up to a few TB per day for modern instruments). Conversely, space-borne instruments can observe below the ionospheric cut-off (10 MHz) and can be placed closer to the studied object.

Several tools have been developed in the last decade for sharing space physcis data. Data visualization tools developed by The CDPP (http://cdpp.eu, Centre de Données de la Physique des Plasmas, in Toulouse, France) and the University of Iowa (Autoplot, http://autoplot.org) are available to display and analyse space physics time series and spectrograms. A planetary radio emission simulation software is developed in LESIA (ExPRES: Exoplanetary and Planetary Radio Emission Simulator). The VESPA (Virtual European Solar and Planetary Access) provides a search interface that allows to discover data of interest for scientific users, and is based on IVOA standards (astronomical International Virtual Observatory Alliance). The University of Iowa also develops Das2server that allows to distribute data with adjustable temporal resolution.

MASER is making use of all these tools and standards to distribute datasets from space and ground radio instruments available from the Observatoire de Paris, the Station de Radioastronomie de Nançay and the CDPP deep archive. These datasets include Cassini/RPWS, STEREO/Waves, WIND/Waves, Ulysses/URAP, ISEE3/SBH, Voyager/PRA, Nançay Decameter Array (Routine, NewRoutine, JunoN), RadioJove archive, swedish Viking mission, Interball/POLRAD... MASER also includes a Python software library for reading raw data.

SPICE FOR ESA PLANETARY MISSIONS. M. Costa[1], ESA/ESAC Camino Bajo del Castillo s/n, Ur. Villafranca del Castillo, 28692 Villanueva de la Canada, Madrid, Spain, marc.costa@esa.int.

Introduction: SPICE is an information system the purpose of which is to provide scientists the observation geometry needed to plan scientific observations and to analyze the data returned from those observations. SPICE is comprised of a suite of data files, usually called kernels, and software -mostly subroutines [1]. A customer incorporates a few of the subroutines into his/her own program that is built to read SPICE data and compute needed geometry parameters for whatever task is at hand. Examples of the geometry parameters typically computed are range or altitude, latitude and longitude, phase, incidence and emission angles, instrument pointing calculations, and reference frame and coordinate system conversions. SPICE is also very adept at time conversions.

The ESA SPICE Service: The ESA SPICE Service (ESS) leads the SPICE operations for ESA missions. The group generates the SPICE Kernel Datasets (SKDs) for missions in operations (ExoMars 2016, Mars Express) missions in development (BepiColombo, JUICE) and legacy missions (Rosetta, Venus Express). ESS is also responsible for the generation of SPICE Kernels for Solar Orbiter. The generation of SKDs includes the operation software to convert ESA orbit, attitude, payload telemetry and spacecraft clock correlation data into the corresponding SPICE format. ESS also provides consultancy and support to the Science Ground Segments of the planetary missions, the Instrument Teams and the science community. ESS works in partnership with NAIF.

Providing the best data: The quality of the data contained on a SKD is paramount. Bad SPICE data can derive to the determination of wrong geometry and wrong geometry can jeopardizes science results. Because of that ESS, in collaboration with NAIF is focused on providing the best SKDs possible. Kernels can be classified as Setup Kernels (Frame Kernels that describe Reference Frames of a given S/C, Instrument Kernels that describe a given sensor FoV and other characteristics, etc.) and Time-varying Kernels (SPK and CK kernels that provide Trajectory and Orientation data, SCLK that provide Time Correlation Data, etc.). Setup Kernels are iterated with the different agents involved in the determination of the data contained in those kernels (Instrument Teams the Science Ground Segment, etc.) and Time-varying kernels are autimatically generated by the ESS SPICE Operational pipeline to feed the Operational kernels that are used in the day-to-day work of the missions in operations (planning and data analysis). These Time-varying kerenels are peer-reviewed a posteriori for the consolidation of SKDs that are archived in the PSA and PDS.

Status of the Kernel Datasets: The current status of the SKDs for the before mentioned missions will be described in this contribution. In general, the ESS is reviewing the legacy and operational datasets and is developing the ones for the future missions. For instance, the Rosetta, Mars-Express and Venus-Express SKDs have just been reviewed and updated.

SPICE Kernels Archived in the PSA. ESS is also responsible for the generation of PDS3 and PDS4 formatted SPICE Archives that are published by the PSA. ESS in close collaboration with NAIF peer-reviews the operational kernels for the PSA [2] to publish being compliant with the Planetary Data System (PDS) standards and uses them in the processes that require geometry computations.

Extended Services: The ESS offers other services beyond the generation and maintenance of SPICE Kernels datasets, such as configuration and instances for WebGeocalc and Cosmographia for the ESA Missions [3].

SPICE-Enhanced Cosmographia. NAIF offers for public use a SPICE-enhanced version of the open source visualization tool named Cosmographia. This is an interactive tool used to produce 3D visualizations of planet ephemerides, sizes and shapes; spacecraft trajectories and orientations; and instrument field-of-views and footprints. ESS Service provides the framework and configuration in order to load the ESA Planetary Missions in Cosmographia, this contribution will demonstrate its usage within the ESA Planetary missions.

WebGeocalc. The WebGeocalc tool (WGC) provides a web-based graphical user interface to many of the observation geometry computations available from the "SPICE" system. A WGC user can perform SPICE computations without the need to write a program; the user need have only a computer with a standard web browser. WGC is provided to the ESS by NAIF. This contribution will outline the WGC instances for ESA Planetary missions.

References: [1] Acton C. (1996) *Planet. And Space Sci., 44,* 65-70. [2] Bessel, S. et al., (2017) *Planet. And Space Sci.* (submitted). [3] Acton, C. et al., (2017) *Planet. And Space Sci.*

SPICE-BASED PYTHON PACKAGES FOR ESA SOLAR SYSTEM EXPLORATION MISSION'S GEOMETRY EXPLOITATION.

M. Costa[1], M. Grass[2,]
[1]ESA/ESAC Camino Bajo del Castillo s/n, Ur. Villafranca del Castillo, 28692 Villanueva de la Canada, Madrid, Spain, marc.costa@esa.int, [2]IRS, Univesity of Stuttgart, Pfaffenwaldring 29, 70569 Stuttgart, Germany.

Introduction: SPICE is an information system the purpose of which is to provide scientists the observation geometry needed to plan scientific observations and to analyze the data returned from those observations. SPICE is comprised of a suite of data files, usually called kernels, and software -mostly subroutines [1]. A customer incorporates a few of the subroutines into his/her own program that is built to read SPICE data and compute needed geometry parameters for whatever task is at hand. Examples of the geometry parameters typically computed are range or altitude, latitude and longitude, phase, incidence and emission angles, instrument pointing calculations, and reference frame and coordinate system conversions. SPICE is also very adept at time conversions.

The ESA SPICE Service: The ESA SPICE Service (ESS) leads the SPICE operations for ESA missions. ESS generates the SPICE Kernel Datasets (SKDs) for missions in operations (ExoMars 2016, Mars Express) missions in development (BepiColombo, JUICE) and legacy missions (Rosetta, Venus Express). ESS is also responsible for the generation of SPICE Kernels for Solar Orbiter. The generation of SKDs includes the operation software to convert ESA orbit, attitude, payload telemetry and spacecraft clock correlation data into the corresponding SPICE format. ESS also provides consultancy and support to the Science Ground Segments of the planetary missions, the Instrument Teams and the science community. ESS works in partnership with NAIF [2]. In addition to the services described in the previous section, ESS is developing several services in the shape of Python packages to enhance the exploitation of SPICE data.

spiops a Python package for SPICE: spiops is a Python package that uses SpiceyPy to use SPICE Toolkit APIS to provide higher-level functions than the ones available with SPICE. spiops is aimed to assists the users to extend the usage of SPICE.

These functions have been identified on the author's and ESA colleagues that use SPICE day-to-day work from having to implement multiple times a series of SPICE APIs to obtain a given derived functionality. Functionalities vary from the computation of the illumination of a given Field-of-View to obtaining the coverage of a given spacecraft (S/C) for a particular meta-kernel, plotting Euler Angles or comparing different kernels. There are three different levels of functions used:

1. SPICE based derived functions
2. non-SPICE based derived functions
3. an object-oriented SPICE interface

The underlying idea of spiops is to be used as a multi-user and multi-disciplinary pool of re-usable SPICE based functions and to provide an easier interface to certain SPICE functionalities with objects to provide cross mission and discipline support of SPICE for ESA Planetary and Heliophysics missions.

spiops is available as a PyPi package and in Bitbucket and is fully documented with ReadTheDocs. This contribution will further describe spiops and will go through several usecases for it.

adcsng a SPICE processing pipeline Python package: SKDs of missions in operations (Mars-Express and ExoMars2016) are regularly updated in missions with new predicted and reconstructed trajectory and attitude information -usually provided by the mission's flight dynamics team- and with Housekeepig Telemertry that provides information of moveable parts of the S/C or the science payload. With these data time-varying kernels are generated (SPKs, CKs and SCLKs) with an automatic processing pipeline: the Auxiliary Data Conversion System next generation (adcsng). adcsng is made available to the ESA missions as a Python Package.

spispy, yet another Python package for SPICE: Although this package is not open to contributions, ESS is also developing a Python package which in combination with adcsng and spiops will provide to the ESA SPICE users a web-based quick-look of a complete SKD. With spispy functionalities, the user will be available to quickly evaluate the state of a given SKD and at the same time to obtain a glimpse of the status of the S/C and the basic geometry of the mission scenario (Measured Attitude, Trajectory, Phase function on a given mission target, Events timeline, etc.) with the kernels generated by adcsng.

References: [1] Acton C. (1996) *Planet. And Space Sci., 44,* 65-70. [2] Costa, M., (2018) this conference.

BRIDGING ARCHIVAL STANDARDS: BUILDING SOFTWARE TO TRANSLATE METADATA BETWEEN PDS3 & PDS4. C. M. De Cesare[1] and J. H. Padams[1], [1]Jet Propulsion Laboratory, California Institute of Technology (4800 Oak Grove Dr., Pasadena, CA 91109-8099, United States of America)

Background: In the PDS3 archiving standard, metadata was labeled using the Object Description Language (ODL). ODL is a proprietary language developed and maintained by the PDS. Variations in the useage of ODL & in the implementation of data products ultimately led to inconsistencies in the implementation of PDS standards, negatively impacting usability & interoperatbility.

In the current PDS4 standard, metadata is labeled using XML (Extensible Markup Language). XML is an international standard that provides a standard syntax and structure for describing data. The adoption of XML allows for increased consistency of metadata, which, in turn, improves accessibility and usability of the archived data.

The need for software tools: Datasets that were archived with the PDS prior to the adoption of PDS4 will retain their PDS3 metadata labels. Over time, certain datasets are being converted into PDS4.

Current & upcoming missions are using PDS4 XML labels to represent their metadata. Some of these missions will continue to produce PDS3 ODL labels alongside the PDS4 XML labels, due to the use of legacy components in their ground data processing systems.

Both of the above situations necessitate translation between PDS3 and PDS4 metadata labels. The translation process requires manual, detail-oriented work, and relies on deep knowledge of PDS4. The goal of translation is to develop a mapping between PDS3 ODL keywords and PDS4 XML XPaths. (An XPath is a reference to a specific location in an XML document.) For more complex datasets, it’s not always possible to achieve a one-to-one mapping from a PDS3 ODL keyword to a PDS4 XML XPath, due to the hierarchical structure of PDS4 labels.

The complexity of the aforementioned translation process calls for a software tool capable of automating this work.

The software solution: The PDS develops & maintains a software application, the Generate Tool[1], that is used to produce a PDS4 XML label using (1) the PDS3 ODL label and (2) a template of the PDS4 label, written for the Apache Velocity[2] templating engine. The Generate Tool extracts metadata values from the PDS3 label and plugs them into the PDS4 label template, dynamically producing a PDS4 label containing the same metadata as the PDS3 original.

The PDS Imaging Node has developed the Label Mapping Tool (LMT), a software tool which leverages the Velocity template in order to produce a list of mappings from PDS3 keyword to PDS4 XPath in a CSV (Comma Separated Value) file. This software automates the comparison of PDS3 to PDS4, and also aids the developer performing the archive conversion by identifying any PDS3 keywords that may have overlooked during development of the PDS4 archive and Velocity template.

Conclusion: The Label Mapping Tool is still in development, but a beta release was completed in early 2018. The PDS Imaging Node plans to release the Label Mapping Tool as open source at https://github.com/nasa-pds.

This presentation will discuss the features and benefits of the software, as well as the challenges and successes encountered during its development.

References: [1] NASA PDS (2017) *Software*. Retrieved from https://pds.nasa.gov/pds4/software.
[2] The Apache Software Foundation. (2016) *Welcome.* Retrieved from http://velocity.apache.org.

3D VISUALIZATION FOR PLANETARY MISSIONS. A. W. DeWolfe,[1] K. Larsen,[1] D. Brain[1], [1]Laboratory for Atmospheric and Space Physics (LASP), University of Colorado, Boulder, CO, alex.dewolfe@lasp.colorado.edu, kristopher.larsen@lasp.colorado.edu, david.brain@lasp.colorado.edu.

We have developed visualization tools for viewing planetary orbiters and science data in 3D for both Earth and Mars, using the Cesium open-source Javascript library. Our tools allow viewers to visualize the position and orientation of one or more spacecraft, as well as science data and atmospheric models. Other 3D viewers exist but are expensive to develop and maintain, and often require the user to download and install standalone software. Cesium and LaTiS are both free and open-source, and allow the user to interact with the 3D visualization and preview science data on the mission's website.

We initially developed this tool as a 3D Mars viewer for NASA's Mars Atmosphere and Volatile EvolutioN (MAVEN)[1] mission. MAVEN has been collecting data at Mars since September 2014, exploring the planet's upper atmosphere, ionosphere, and interactions with the Sun and solar wind. We used Cesium to display MAVEN's orbit around Mars, its orientation as it points in different directions during its orbit, and the science data it collects from the Martian atmosphere and the solar wind. Some of the instruments collect data along the path of the spacecraft ("in situ" measurements), and some measure magnetic fields or particle velocities that need to be displayed as 3D vectors branching out from the orbital path. Viewers can choose to display the spacecraft in the planet's reference frame so that Mars remains stationary, or an inertial reference frame to watch the planet rotate. We have also added the capability to display the M-GITM atmospheric models overlaid on the surface of Mars.

In order to stream the data from a set of ASCII files stored on our server, we use LaTiS[2], a free, open-source unified data access service developed here at LASP to stream time-series data and convert between formats. This allows us to stream the science data into a JSON object that can be easily displayed by the Cesium code. In addition, LaTiS can use NASA's SPICE libraries to compute ephemeris information and provide it as JSON, so that the ephemeris and pointing information for any spacecraft can be displayed in Cesium.

After adapting Cesium and LaTiS to display MAVEN at Mars, we began modifying it to display NASA's Magnetospheric MultiScale (MMS)[3] mission, a four-spacecraft constellation currently orbiting the Earth to measure Earth's magnetic field and its interaction with the solar wind. MMS's orbit and constellation configuration are challenging to visualize, making an interactive 3D viewer highly useful for the science team as well as a useful educational tool.

References:

[1] http://lasp.colorado.edu/home/maven/
[2] https://github.com/latis-data/latis
[3] https://mms.gsfc.nasa.gov/

VESPA: ENLARGING THE VIRTUAL OBSERVATORY TO PLANETARY SCIENCE. S. Erard[1], B. Cecconi[1], P. Le Sidaner[2], A. P. Rossi[3], T. Capria[4], B. Schmitt[5], N. André[6], A. C. Vandaele[7], M. Scherf[8], R. Hueso[9], A. Määttänen[10], B. Carry[11], N. Achilleos[12], C. Marmo[13], O. Santolik[14], K. Benson[12], P. Fernique[15], [1]LESIA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, UPMC Univ. Paris 06, Univ. Paris Diderot, Sorbonne Paris Cité, Meudon, France, [2]DIO-VO/UMS2201 Observatoire de Paris/CNRS, Fr, [3]Jacobs University, Bremen, Ge [4]INAF/IAPS, Rome, It [5]IPAG UGA/CNRS, Grenoble, Fr [6]IRAP/CNRS, Toulouse, Fr [7]IASB/BIRA, Brussels, Be [8]OeAW, Graz, Aut [9]UPV/EHU, Bilbao, Sp [10]LATMOS/CNRS, Guyancourt, Fr [11]OCA, Nice & IMCCE/Obs. Paris/CNRS, Fr [12]University College London, UK [13]GEOPS/CNRS/U. Paris-Sud, Fr [14]IAP, Prague, Cz R. [15]Observatoire de Strasbourg/UMR 7550, Fr

Introduction: Modern space borne instruments produce huge datasets, especially on long-lived missions. This calls for new ways to handle the data, not only to perform mass processing, but also more basically to access them easily and efficiently. Virtual Observatory (VO) techniques developed in Astronomy during the past 15 years can be adapted to address this problem, provided they are enlarged to handle specificities of Solar System studies. The VESPA data access system focuses on applying VO techniques and tools to Planetary Science data, and supports all aspects of Solar System science [1]. VESPA (Virtual European Solar and Planetary Access) is developed in the framework of the EU-funded Europlanet-2020 program started Sept 1st, 2015 for 4 years. The objective of this activity is to facilitate searches in big archives as well as in sparse databases, to provide simple data access and on-line visualization tools, and to allow small data providers to make their data available in an interoperable environment with minimum effort. This system makes intensive use of studies and developments led in Astronomy (International Virtual Observatory Alliance, IVOA [2]), Solar Physics (HELIO), and space archive services (International Planetary Data Alliance, IPDA [3]).

Data services: the VESPA architecture [1] is based on a new data access protocol, a specific user interface to query the available services, and intensive usage of tools and standards developed for the Astronomy VO [4]. The Europlanet data access protocol, EPN-TAP, relies on the general TAP (Table Access Protocol) mechanism associated to a set of parameters that describe the contents of a data service [1][5]. These parameters are defined to enable queries on quantities relevant to the scientific user, including observational and instrumental conditions. Data services are required to return answers formatted as VOtables, which are handled by all standard VO tools.

Data services are installed at their respective provider institutes and are declared in the standard IVOA registries, so that they are always visible and reachable from query interfaces. At the time of writting, 37 data services are publicly open, and about 20 more are being finalized (see, e. g., [6]). They encompass a wide scope, including surfaces, atmospheres, magnetospheres and plasmas, small bodies, experimental data such as spectroscopy in solid phase, heliophysics, exoplanets, and selected amateur data services. ESA's Planetary Science Archive (PSA) will get an EPN-TAP interface in 2018, and bridges with PDS4 are being studied.

Data access: Although accessible in many ways, EPN-TAP data services are best queried from the optimized VESPA user interface, or portal (Fig. 1)

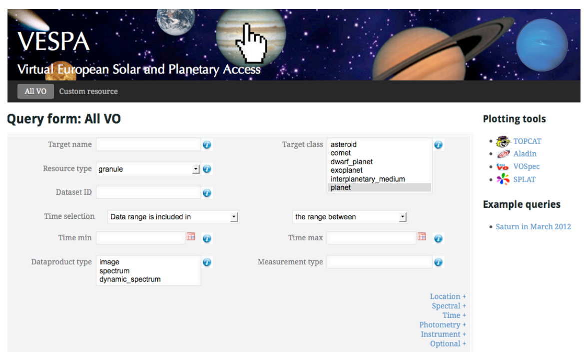


Fig 1: The VESPA search interface: http://vespa.obspm.fr

In the frame of TAP, all data services present a list of granules (usually data files) described by a series of parameters. The Europlanet data access protocol, EPN-TAP, defines a set of mandatory parameters introducing metadata that describe all granules; this is similar to the ObsTAP protocol from IVOA, which describes observational datasets in Astronomy. EPN-TAP metadata introduce both observational and instrumental conditions and are defined to handle the specific diversity and complexity of Planetary Science: ranges on several axes (spatial, temporal, spectral, photometric), measurement type, origin of data, and various references. Location is provided in the most appropriate coordinate system (e.g., sky or planetary coordinates); target-related time (local time and season, through Ls) can be provided when relevant. The VESPA portal uses the mandatory parameters to search for individual granules in all registered data services at once, allowing for discovery of data content unknown to the user. In addition, specific parameters may also be used to describe individual services in more details, and can be used to identify granules more precisely when querying a single service.

All granules provide a link to a data file, or include the data itself in the table when possible (e.g., for scalar quantities). Data description parameters are used to identify adequate VO tools to access, plot and handle the data correctly. They not only provide a description of the file format, but also specify the dimensions, units, and physical quantities, relying on IVOA data models extended for VESPA. For instance, spectra and images are handled in different tools, and spectra measured in radiance or in reflectance are handled differently by the spectral tools.

Tools: Metadata are smoothly transferred from the VESPA portal to VO tools according to the IVOA SAMP protocol, with no user intervention. Standard VO tools are connected to the VESPA portal so that they readily display metadata, e. g., spatial footprints are plotted on a 3D sphere in Aladin or Mizar; other metadata such as local time, instrument modes, etc… can be plotted in 2D or 3D with TOPCAT.

The data themselves can be transferred in a similar way for display and standard analyses. Data description is used to select appropriate tools, e.g. TOPCAT will handle all types of tables, Aladin most images and spectral cubes, CASSIS and SPLAT-VO spectra in general (Fig. 2), 3Dview can plot measurements along a spacecraft trajectory, Autoplot is dedicated to extracting data from long time series, etc. Both Aladin and TOPCAT can produce multiresolution maps from large datasets.

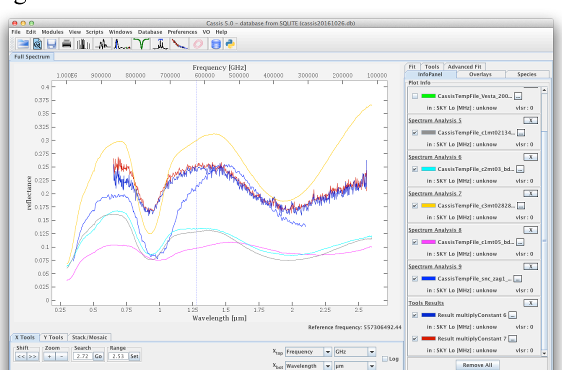


Fig 2: telescopic spectra of Vesta compared to basaltic meteorites from the PDS spectral library in CASSIS.

Most of these tools have been updated to support Planetary Science and to handle specificities of Solar System data, e. g., measurements in reflected light, coordinate systems on surfaces and in magnetospheres, etc. Other, non-VO tools are provided with a SAMP interface so that they can be included in pipe-lines (e. g., ImageJ which now provides conversions and image processing functions to the VO) and in some case new applications have been developed for VESPA (e. g., to handle fits images of planetary surfaces [7], or PDS3 spectral cubes). A significant on-going activity is the development of a connection between the VO world and Geographic Information Systems (GIS), so that EPN-TAP data services can provide links to WMS services which are then handled in QGIS; the intermediate VO layer allows for powerful search functions in the data, and cross-examinations with other datasets.

Computation services: another goal is to connect on-line computation services with interfaces similar to the data services, so as to compare observations and simulations more routinely. This activity has obvious applications, e. g., for radiative transfer in planetary atmospheres. A complementary aspect is to provide low level computation functions on-line, e.g., averages, resampling, deconvolution, etc. This is currently supported only to some extent by standard VO tools and ImageJ; in addition, higher level processing such as retrieval of Hapke parameters, multivariate analyses, etc, would also be beneficial and are being studied.

External contributions: A procedure has been identified to install data services with little resources, and hands-on sessions are organized twice a year at EGU and EPSC conferences in Europe (see VESPA web site). This is expected to favor the installation of services by individual research teams, e. g. to distribute derived data related to a published study. In complement, regular discussions are held with big data providers, starting with space agencies (IPDA). In parallel, a Solar System Interest Group has just been started in the IVOA, where several VESPA partners contribute; the goal is here to adapt existing astronomy standards to Planetary Science.

Acknowledgements:
The Europlanet 2020 Research Infrastructure project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654208.
VESPA web site: http://www.europlanet-vespa.eu

References:
[1] Erard et al 2017, *PSS* 10.1016/j.pss.2017.05.013. ArXiv 1705.09727 [2] http://www.ivoa.net [3] https://planetarydata.org [4] Erard et al 2014 *Astron & Comput* **7-8**, 52-61. ArXiv 1407.5738 [5] Erard et al 2014 *Astron & Comput* **7-8**, 52-61. ArXiv 1407.5738 [6] Giardino et al 2018 *this conference* [7] Marmo et al 2018 *this conference*

Taming Pipelines, Users, and High Performance Computing with Rector. N. M. Estes, K. S. Bowley, K. N. Paris, V. H. Silva, M. S. Robinson, School of Earth and Space Exploration, Arizona State University

Introduction: The Lunar Reconnaissance Orbiter Camera (LROC) Science Operations Center (SOC) receives and processes ~450 gigabits of data every day. These data get uncompressed and processed into engineering data records (EDRs) and calibrated data records (CDRs). During processing, there are many steps to catalog, calibrate, calculate geometry, validate end products, and many other steps (132 pipeline procedures in all), and it requires a well-designed system to orchestrate all of these steps for thousands of products each week. In addition to this baseline processing, users require a wide variety of tasks to be run on thousands of images for map projecting, mosaicking, photometric correction, and other tasks. This processing is currently coordinated by Rector over a 14-node processing cluster running a total of 634 CPU cores.

The baseline requirements for the job management system included the ability to automatically allocate CPUs and RAM on a processing node, coordinate between hundreds of CPU cores over many nodes, provide a GUI interface for operations staff to monitor processing and handle exceptions, and provide a way for users to distribute arbitrary jobs across the cluster without specialized knowledge. Before creating Rector, the LROC SOC evaluated several job control systems including HiRISE Conductor [1], Condor [2], and Torque [3]. While these options handled some of the requirements, no single solution met all requirements. After developing an initial prototype, the team decided to develop a job control system that met all of the needs of the LROC SOC.

Job Management: Rector handles two categories of jobs: automated pipeline jobs and user jobs. These two categories are handled differently to meet the needs of the users of each job type. In both cases, the job queues are managed via a central PostgreSQL database using row exclusive locks to ensure that each job runs only once. Logs of every job run by Rector are also kept in the database with all information necessary to re-run the job in the future if necessary.

The GUI interface for LROC SOC operations staff is written using the Ruby on Rails framework. The Rector daemon that runs on each processing node, command line tools for job management, and administrative tools are written in Ruby.

Pipeline Job Management: Pipeline jobs are automated procedures that ingest all files received from the mission operation center and perform all necessary processing steps to generate EDRs, CDRs, browse products, histograms; as well as provide quality control, statistics calculation, and many other tasks necessary for LROC operations. These jobs are managed through a GUI interface (Figure 1) where the LROC SOC operations staff can see all job status information and handle any exceptions that may have occurred. This interface provides a general overview of every node in the cluster and every configured pipeline procedure, along with the ability to drill down to the fine details if necessary (Figures 2, 3).

The pipeline procedure configuration allows for procedures to be prioritized (for example, importing SPICE kernels is a higher priority than producing CDR products). Operations staff can further control how things are run by specifying a regular expression that gets matched against the procedure name to control which procedures are allowed to run on a given node.

Details on how the LROC pipelines work can be found in the "Scalable Data Processing with the LROC Processing Pipelines" abstract also submitted to this conference.

User Job Management & Security: User jobs are handled differently than pipeline jobs in Rector to allow arbitrary commands to be submitted to the cluster. The command line tools for managing user jobs were modeled after the high performance computing standard OpenPBS command line tools [4]. User jobs run at a lower priority than pipeline jobs, and when running user jobs, Rector attempts to balance jobs submitted by different users as much as possible. For example, if user A is using all available CPUs in the cluster and user B submits a set of jobs, as jobs finish, Rector

Procedure	Total	Total Incomplete	Running	Queued	Paused	Total Completed	Succeeded	Timed Out	Failed	Aborted	Orphaned	Unhandled Exceptions
lroc_edr_p10	0	0	0	0	0	0	0	0	0	0	0	0
lroc_edr_p20	1	0	0	0	0	1	1	0	0	0	0	0
lroc_edr_p30	3	0	0	0	0	3	3	0	0	0	0	0
lroc_fdf13_p10	1	0	0	0	0	1	1	0	0	0	0	0
lroc_fdf14_p10	1	0	0	0	0	1	1	0	0	0	0	0
lroc_fdf21_p10	1	0	0	0	0	1	1	0	0	0	0	0
lroc_fdf22_p10	1	0	0	0	0	1	1	0	0	0	0	0
lroc_fdf29_p10	1	0	0	0	0	1	1	0	0	0	0	0
lroc_fdf30_p10	1	0	0	0	0	1	1	0	0	0	0	0
lroc_fdf30_p20	1	0	0	0	0	1	1	0	0	0	0	0
lroc_genmk_p10	12	1	1	0	0	11	9	2	0	0	0	0
lroc_geo_p10	576	0	0	0	0	576	294	0	282	0	0	82

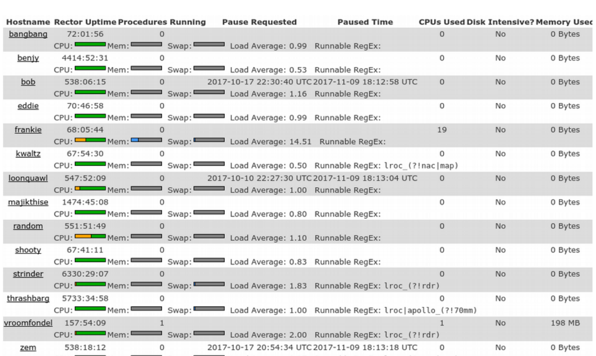
Figure 1: Small selection of pipeline procedures showing current procedure queue status as well as log status over a user-defined time interval.

will start user B's jobs in preference until the number of CPUs used by user A and user B match, at which time it will alternate between starting jobs from user A and user B to maintain that balance. Likewise, if user C also submits jobs, user C's jobs will be selected in preference until a balance between all three users is achieved. Rector will never stop a running job. Prioritization and user-balancing only occur when Rector selects a new job to run.

For security purposes, user jobs are run as the user that submitted the job. This way the user's normal permissions and access restrictions also apply to the jobs they submit to Rector. Because these jobs are queued in a central database and the user needs to be able to access the database to add jobs to the queue, a security mechanism is in place to ensure that users can only submit jobs that run as themselves and that submitted commands cannot be modified. To achieve this, each user has a public/private key-pair for Rector. This private key should be protected the same as any SSH, PGP, or other private key. When submitting a job, the job submission tool generates a cryptographic signature of the command to be run using the user's private key, and this signature is stored in the database. When Rector selects a user job to run, it loads the user's public key and verifies that the signature matches the command to run. After the initial setup of the key, this process is completely transparent to the user, but if a signature does not match a command, Rector will generate an error and refuse to run the job.

Results: To date, Rector has shepherded over 4.3 million EDR and CDR products through the LROC pipeline. Between those products, and all the other pipeline procedures, Rector has run over 71.2 million pipeline jobs since 2009. In addition to the pipeline jobs, Rector has also been used to handle more than 40.1 million user jobs in that time.

Rector has proven to be a reliable and easy-to-use job management tool at the LROC SOC. It has grown from an initial processing cluster of 6 nodes and 12 CPU cores to the current cluster size of 14 nodes and



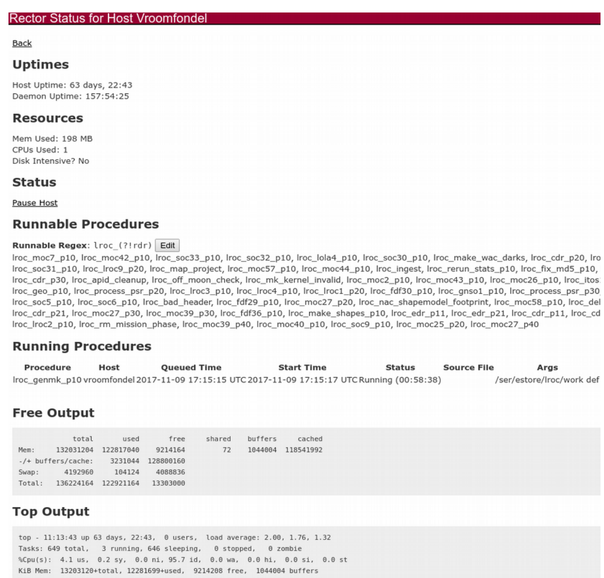
Hostname	Rector Uptime	Procedures Running	Pause Requested	Paused Time	CPUs Used	Disk Intensive?	Memory Used
bangbang	72:01:56	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 0.99 Runnable RegEx:						
benjy	4414:52:31	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 0.53 Runnable RegEx:						
bob	538:06:15	0	2017-10-17 22:30:40 UTC	2017-11-09 18:12:58 UTC	0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 1.16 Runnable RegEx:						
eddie	70:46:58	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 0.99 Runnable RegEx:						
frankie	68:05:44	0			19	No	0 Bytes
	CPU: Mem: Swap: Load Average: 14.51 Runnable RegEx:						
kwaltz	67:54:30	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 0.50 Runnable RegEx: lroc_(?!nac|map)						
loonquawl	547:52:09	0	2017-10-10 22:27:30 UTC	2017-11-09 18:13:04 UTC	0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 1.00 Runnable RegEx:						
majikthise	1474:45:08	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 0.80 Runnable RegEx:						
random	551:51:49	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 1.10 Runnable RegEx:						
shooty	67:41:11	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 0.83 Runnable RegEx:						
strinder	6330:29:07	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 1.83 Runnable RegEx: lroc_(?!rdr)						
thrashbarg	5733:34:58	0			0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 1.00 Runnable RegEx: lroc|apollo_(?!70mm)						
vroomfondel	157:54:09	1			1	No	198 MB
	CPU: Mem: Swap: Load Average: 2.00 Runnable RegEx: lroc_(?!rdr)						
zem	538:18:12	0	2017-10-17 20:54:34 UTC	2017-11-09 18:13:18 UTC	0	No	0 Bytes
	CPU: Mem: Swap: Load Average: 0.97 Runnable RegEx: lroc_(?!nac|map)						

Figure 2: List of nodes in the processing cluster with summary information on procedures running, CPU cores used, and other diagnostic information.

634 CPU cores. Rector's prioritization capabilities allow the LROC SOC operations staff to manage how jobs are run when necessary, but most of the time, Rector's own ability to manage resources handles all job marshaling in a completely hands-off manner. The job logging allows operations staff to look back on processing history at any time, and job errors are presented to operations staff in an intuitive way that allows human intervention when necessary.

Future Work: Rector currently has a limitation when hundreds of short-duration jobs are submitted to a cluster with a large number of CPU cores. Rector will successfully run the jobs, but it will not use all available cores in that case due to the timing of the exclusive lock necessary in the database to ensure that the Rector daemons coordinate successfully between themselves. To avoid this issue, it is recommended that users avoid submitting jobs that run in less than 30 seconds. It is suggested that these short-duration jobs be batched-up in chunks that take longer than 30 seconds, so that Rector is able to keep all cores in the cluster running at full capacity. Work to eliminate that inconvenience would be helpful given enough time in the development schedule.

References: [1] Schaller, C.~J.\ 2006, 37th Annual Lunar and Planetary Science Conference, 37 [2] Douglas Thain, Todd Tannenbaum, and Miron Livny, "Distributed Computing in Practice: The Condor Experience" Concurrency and Computation: Practice and Experience, Vol. 17, No. 2-4, pages 323-356, February-April, 2005. [3] Torque http://www.adaptivecomputing.com/products/open-source/torque/ [4] OpenPBS http://www.mcs.anl.gov/research/projects/openpbs/



Rector Status for Host Vroomfondel

Back

Uptimes

Host Uptime: 63 days, 22:43
Daemon Uptime: 157:54:25

Resources

Mem Used: 198 MB
CPUs Used: 1
Disk Intensive? No

Status

Pause Host

Runnable Procedures

Runnable Regex: lroc_(?!rdr) Edit
lroc_moc7_p10, lroc_moc42_p10, lroc_soc33_p10, lroc_soc32_p10, lroc_lola4_p10, lroc_soc30_p10, lroc_make_wac_darks, lroc_cdr_p20, lro
lroc_soc31_p10, lroc_lroc9_p20, lroc_map_project, lroc_moc57_p10, lroc_moc44_p10, lroc_ingest, lroc_rerun_stats_p10, lroc_fix_md5_p10,
lroc_cdr_p30, lroc_apid_cleanup, lroc_off_moon_check, lroc_mk_kernel_invalid, lroc_moc2_p10, lroc_moc43_p10, lroc_moc26_p10, lroc_itos
lroc_geo_p10, lroc_process_psr_p20, lroc_lroc3_p10, lroc_lroc4_p10, lroc_lroc1_p20, lroc_fdf30_p10, lroc_gnso1_p10, lroc_process_psr_p30
lroc_soc5_p10, lroc_soc6_p10, lroc_bad_header, lroc_fdf29_p10, lroc_moc27_p20, lroc_nac_shapemodel_footprint, lroc_moc58_p10, lroc_del
lroc_cdr_p21, lroc_moc27_p30, lroc_moc39_p30, lroc_fdf36_p10, lroc_make_shapes_p10, lroc_edr_p11, lroc_edr_p21, lroc_cdr_p11, lroc_cd
lroc_lroc2_p10, lroc_rm_mission_phase, lroc_moc39_p40, lroc_moc40_p10, lroc_soc9_p10, lroc_moc25_p20, lroc_moc27_p40

Running Procedures

Procedure	Host	Queued Time	Start Time	Status	Source File	Args
lroc_genmk_p10	vroomfondel	2017-11-09 17:15:15 UTC	2017-11-09 17:15:17 UTC	Running (00:58:38)		/ser/estore/lroc/work def

Free Output

```
             total       used       free     shared    buffers     cached
Mem:     132031204  122817040    9214164         72    1044004  118541992
-/+ buffers/cache:    3231044  128800160
Swap:      4192960     104124    4088836
Total:   136224164  122921164   13303000
```

Top Output

```
top - 11:13:43 up 63 days, 22:43,  0 users,  load average: 2.00, 1.76, 1.32
Tasks: 649 total,   3 running, 646 sleeping,   0 stopped,   0 zombie
%Cpu(s):  4.1 us,  0.2 sy,  0.0 ni, 95.7 id,  0.0 wa,  0.0 hi,  0.0 si,  0.0 st
KiB Mem:  13203120+total, 12281699+used,  9214208 free,  1044004 buffers
KiB Swap:  4192960 total,   104124 used,  4088836 free. 11854199+cached Mem
```

Figure 3: Node detail page showing a variety of Rector daemon and host diagnostic information.

Standards-Based Open-Source Planetary Map Server: Lunaserv N. M. Estes, V. H. Silva, K. S. Bowley, K. K. Lanjewar, M. S. Robinson, School of Earth and Space Exploration, Arizona State University, nme@ser.asu.edu

Introduction: The Lunar Reconnaissance Orbiter Camera (LROC) is operated from the Science Operations Center (SOC) on the ASU campus. SOC operations require dynamically generated maps to support the functionality of a wide variety of tools and applications. Some of these uses include JMoon [1], web sites, and Geographic Information System (GIS) software. In addition to these applications, specialized map requests are used to generate frames for videos, figures for documents and publications, production of outreach materials, etc. To meet the widest variety of uses possible with a map tool and single set of map data, we first investigated the Web Map Service (WMS) standard. At the time, there were no WMS packages that supported non-Earth coordinate reference system (CRS) definitions out of the box, and many solutions had limitations involving large global sets comparable to LROC. As a result, we developed Lunaserv (2009) to provide a WMS compatible server software supporting planetary CRS [2], and capable of serving voluminous (terabytes) data sets.

While the initial version of Lunaserv supported only a small selection of CRS definitions for the Moon, subsequent updates expanded Lunaserv capabilities. Lunaserv now supports many planetary CRS definitions without any additional user configuration. The ability to serve map data for any planetary body in a variety of projections to WMS client software is valuable to many planetary researchers, so in 2013, Lunaserv was released as open-source software [3].

Capabilities: Lunaserv implements the Open Geospatial Consortium (OGC) WMS standard. The WMS standard was chosen because it is a protocol widely used by a variety of GIS software including QGIS, ArcGIS, Grass, OpenLayers, Leaflet, and JMARS. With the WMS standard, Lunaserv provides map data for the largest possible set of GIS data users from a single set of source data [4].

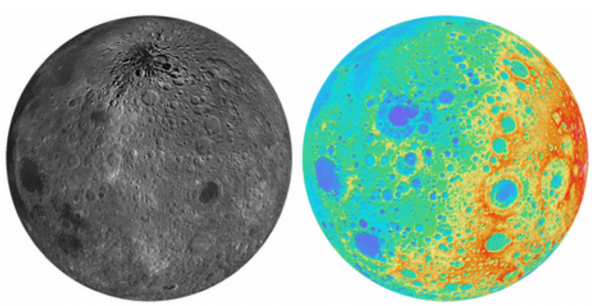


Figure 1: The Moon rendered in an orthographic projection centered at 45° N, 120° E. The left is a LROC WAC global mosaic, and the right is a color shade based on the GLD100 [12].

The WMS standard allows map data to be rendered in a variety of raster formats, and in any CRS understood by the WMS server. A WMS CRS specifies the combination of projection and planetary body spheroid [5]. While the WMS specification recognizes only Earth-based CRS definitions, Lunaserv supports all of the 8,250 IAU2000/IAU2009 planetary CRS definitions, and any arbitrary CRS that can be defined using the proj.4 library [6].

Supported Layer Types: The WMS protocol is primarily focused on presenting 8-bit visual map data for use in web applications or GIS software. Accordingly Lunaserv generates map products from both raster and vector source data..

Raster Image Data (Fig 1): Lunaserv supports both 8-bit grayscale and 24-bit RGB image data. This type of map is loaded from pyramidal TIFFs (PTIFF). These PTIFFs can either have embedded geographic meta-data, or the geographic meta-data can be specified in a separate file. The PTIFFs can also have a 1-bit mask file to specify the area of interest within the PTIFF that should be rendered. The PTIFF filenames for a layer are listed in the layer configuration file, or the list can be loaded from a database.

Lunaserv additionally can serve 32-bit floating point data for compatible client software that understands either 32-bit TIFF or 32-bit VICAR (currently support for Jmoon/Jmars). The source of the high-precision data is an Integrated Software for Imagers and Spectrometers (ISIS) cube [8]. Multiple ISIS cubes of different resolutions can be provided (similar to the internal structure of a pyramidal TIFF), and Lunaserv will render each request using the ISIS cube that is the most appropriate resolution for the map request. Future work may include support for 32-bit TIFF format .

Vector Data (Fig 2): Lunaserv loads vector data from flat files, shapefiles, or a database in the form of points, line-strings, polygons, annotations, or grids. The grid vector type is a specialized layer

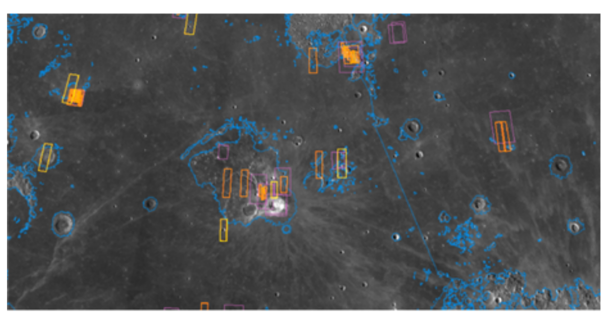


Figure 2: WAC global mosaic with ROI, DTM, Anaglyph, and shapefile RDR product layers overlaid as examples of vector layers.

configuration that requires no input data, but can generate any size latitude/longitude grid specified in the layer configuration file.

Lunaserv supports the PostgreSQL database by default, but support for other databases is possible by creating a simple software plugin. All database operations support a rich set of filtering capabilities and can use a predefined set of 5° bins to limit the query results to the area of interest for performance reasons.

Illumination (Fig 3): Lunaserv can generate illuminated maps using simple day/night shading, or topography-based illumination. Both options render the requested illumination dynamically based on the sub-solar point calculated using the NAIF SPICE toolkit [7]. For topographic-based illumination, a Digital Terrain Model (DTM) in the form of a 32-bit ISIS cube file is used to provide the necessary elevation data [8].

The illumination types use the current time by default, or will render the illumination conditions for any provided time or sub-solar point. As with the 32-bit raster image data support, work may be done in the future to support the PTIFF format.

Usage: The LROC SOC uses Lunaserv to provide data for operations, data portals, web site context maps, PDS web interface, “Where is LRO” webpage, digitizing, video generation, and other activities [9]. In addition to SOC uses, outside researchers also use Lunaserv for visualization with a variety of GIS software packages. The public Lunaserv hosted by the LROC SOC contains all of the LROC map projected PDS products. For demonstration purposes, it additionally serves base imagery, illumination and nomenclature for Mercury, Venus, Earth, Mars, Io, Ganymede, Europe, Callisto, Rhea, Tethys, Iapetus, Dione, Enceladus, Pluto, Charon, Vesta, and Ceres. Based on log file analysis, Lunaserv has been used by other researchers with QGIS, ArcGIS, Google Earth, OpenSceneGraph, OpenLayers, and Leaflet [10]. On average, the public Lunaserv service hosted by the LROC SOC handles more than 20k map requests per day, and during periods of high activity, has handled over 600k map requests in a single day.

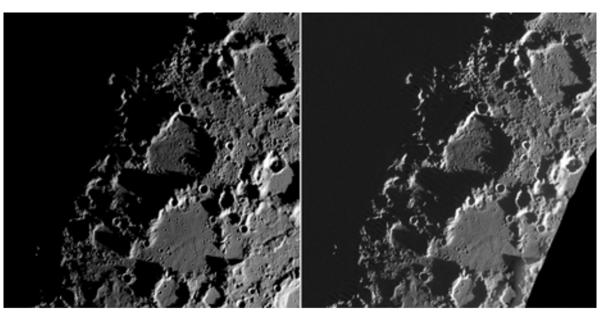


Figure 3: The north pole of the Moon on 2012-357. The left is a synthetic illumination map rendered by Lunaserv using the GLD100 DEM [12]. The right is a composite of actual LROC WAC observations from 2012-357.

In addition to using the LROC SOC hosted Lunaserv server, Lunaserv can also be installed and used by other groups to host their own map data. One such installation is the Lunaserv service supporting the I4 tool provided by Johnson Space Center [11].

Conclusion: The WMS protocol allows for GIS software users to easily combine data from multiple sources without first downloading or processing the data in any way. Lunaserv leverages this capability and extends it to provide support for the IAU2000/ IAU2009 planetary CRS definitions, and provides support for large global data sets. By making data available using Lunaserv, research groups can make accessing their data faster and easier using software that many GIS users are already familiar with, and exposes the underlying data to uses not originally envisioned without custom development.

References: [1] Christensen, P.R., et. al., JMARS – A Planetary GIS, http://adsabs.harvard.edu/abs/ 2009AGUFMIN22A..06C.

[2] Hare, T. et. al., (2006), LPSC XXXVII, abs. 1931.

[3] Estes, N.M.; et. al.; (2013) Lunaserv Web Map Service: History, Implementation Details, Development, and Uses, http://www.lpi.usra.edu/ meetings/lpsc2013/pdf/2609.pdf.

[4] Dobinson, E., et. al., (2006), LPSC XXXVII, abs. 1463

[5] OGC WMS Standards, http:// www.opengeospatial.org/standards/wms

[6] proj.4 https://trac.osgeo.org/proj/

[7] NAIF Spice Toolkit http://naif.jpl.nasa.gov/naif/ toolkit.html

[8] Hanger, C. D.; et. al.; DEM-based Illumination Simulation in a Web Map Service using Lunaserv; http://lunarscience.nasa.gov/wp-content/uploads/ LSF13P/ Hanger_nlsf2013.pdf.

[9] Estes, N.M.; et al. (2014):, Lunaserv 3 Development and Usage Over the Past Year, http:// www.lpi.usra.edu/meetings/lpsc2014/pdf/2180.pdf

[10] Estes, N.M.; et al. (2015), Lunaserv WMS Software Enhancements, http://www.lpi.usra.edu/ meetings/lpsc2015/pdf/1507.pdf

[11] Stefanov, W. L, et. al. (2015), The I4 Online Query Tool for Earth Observations Data, JSC-CN-33076

[12] Scholten, F., et. al. (2012), , J.GR., 117, E00H17, doi:10.1029/2011JE003926.

THE IMPORTANCE OF GEODETICALLY CONTROLLED DATA SETS: THEMIS CONTROLLED MOSIACS OF MARS, A CASE STUDY. R. L. Fergason[1] and L. Weller[1], Astrogeology Science Center, [1]U.S. Geological Survey, 2255 N. Gemini Drive, Flagstaff, AZ 86001, rfergason@usgs.gov.

Introduction: Geodetically controlled products at global scales are foundational data products [1] that provide a common reference system to enable the accurate co-registration of multiple data sets. Accurate registration is necessary for the precision science required to answer questions that cross-cut disciplines and are potentially key to understanding fundamental questions about our universe. To provide such a foundational product for Mars, we have geodetically controlled and mosaicked Thermal Emission Imaging System (THEMIS) [2] daytime infrared (IR) and nighttime IR images resulting in improved camera pointing and spacecraft position knowledge. The results of this work are kernel files describing these improvements for each image in the control network and controlled, orthoprojected daytime IR and nighttime IR mosaics of Mars at 100 m/pixel scale for the ±65˚ latitude region of Mars, and constitutes a foundational data product for Mars. These mosaics, and the associated network, have improved the registration of the THEMIS IR data set, enhance our knowledge (position, precision, and accuracy) of image placement and the location of small-scale surface features, and provide for improved targeting for current and future orbital acquisition of data and spacecraft landings and planning of spacecraft surface operations.

Definitions and Background: Controlled data are defined as data that have been precisely co-registered relative to one another and to some datum (e.g., altimetry data in planetary science), such that rigorous accuracy and precision information is quantifiable (most often using photogrammetric bundle adjustment techniques). This is in contrast to oft utilized "rubber sheeting" techniques that can result in aesthetically pleasing image mosaics, but lack rigorous statistical information that must be available to quantify the image adjustment. Without rigorous accuracy information, the propagation of error to additional co-registered (e.g., fused) data sets is impossible, making scientific assessments that require accurate positional information problematic. A key output of photogrammetric control is adjusted image pointing knowledge that can be used to generate improved spacecraft position and instrument pointing information (i.e., CK and SPK kernels; SPICE). This updated SPICE can be used by other researchers to properly geometrically locate their data, but meet their specific data processing needs.

Global and regional control networks and mosaics, and the products that result from the improved pointing and position information, are critical for a broad range of applications, such as precision landing on a planetary body and surface change detection studies. An accurate base at the highest resolution reasonable is necessary to facilitate the co-location of data (e.g., data fusion) with known, quantifiable error, and controlled data products provide such a foundation to which all other data sets intersecting the network can then be registered. Controlled products helps meet a common need among researchers to accurately co-locate data sets of varying spatial scales and data types to aide in addressing multi-disciplinary questions and allows for precision science to be accomplished.

A common practice of simply co-registering images to one another (automatic or manually) without using a controlled base to tie to, and retaining the accuracy of that registration, can only be considered a "semi-controlled" product. This process may produce a product that is "good enough" for many science investigations, and is certainly less costly to produce, but the resulting product does not achieve sufficient accuracy, and more importantly the knowledge of that accuracy, to be a controlled product and its value is significantly reduced. As described above, without rigorous accuracy assessments, analyses that depend upon positional accuracy or geometric relationships can be incorrect when hypotheses are tested within the error thresholds.

Methods: The current accepted ground data source for Mars is Mars Orbiter Laser Altimetry (MOLA [3-4]) digital elevation model. The horizontal resolution of this product is 463 m/pixel and the overall horizontal accuracy is ~100 meters [5], which is insufficient to confidently register high-resolution images, such as High Resolution Imaging Science Experiment [6] (spatial scale of 30 m/pixel), to this base due to large differences in spatial resolution. As a global data set of intermediate spatial scale, we have geodetically controlled and mosaicked THEMIS daytime IR and nighttime IR images to enable the accurate co-registration of martian data sets. The THEMIS instrument [2] has attained near global coverage of Mars in the daytime and nighttime IR at a scale of ~100 m/pixel, providing the needed images to geodetically control (i.e., precisely and accurately register in a consistent solution with estimates of uncertainty) these data into a common reference coordinate frame at the sub-pixel level.

We generated the THEMIS IR controlled mosaics using automatic sub-pixel registration (with human oversight) and bundle adjustment software in ISIS3 [7]. We then perform a least squares bundle adjustment of control point image measurements and generate the control networks. We also solve for the formal uncertainties for the exterior ordination of the images (i.e., pointing of the spacecraft and position of each image), the latitude/longitude/radius uncertainties on each point in the control network, and uncertainties in the constrained ground points.

After image-to-image ties are completed, we then tied the THEMIS tiles to ground using an improved Viking MDIM 2.1 network [8]. To take advantage of the high accuracy High/Super Resolution Stereo Colour Imager (HRSC) [9] data and the geometric strength of the global Viking MDIM 2.1 [10-11], we reprocessed the original MDIM 2.1 network incorporating available HRSC level 4 data (which have been well controlled to the MOLA reference frame [9]) as additional ground control. Error propagation showed that 80% (~2700 points) of the final enhanced MDIM 2.1 solution tie points have horizontal accuracies better than 200 meters [8]. This methodology results in a control network and an orthorectified product that has broad applicability. In addition to mosaics, updated camera pointing and spacecraft position kernels of all THEMIS images included in the control networks have been generated. We will deliver final kernels to the Navigation and Ancillary Information Facility (NAIF) so they can include it on the PDS NAIF FTP site.

Results: We have found that errors in image position at the 2-4 pixel level (but as large as 30 pixels) are apparent in uncontrolled data downloaded from the PDS (Figure 1). These errors are primarily due to uncertainties in the THEMIS image start time. This uncertainty is random, and there are no future plans to improve the THEMIS IR camera model further. Controlling the pointing has enabled the correction of these errors and improve both the registration between images and registration to a known coordinate reference frame (i.e., MOLA) at known levels of precision and accuracy. In all tiles, the accuracy of image position is less than a single pixel, and the 3-sigma residual is also less than a single pixel. The position of a single THEMIS image is commonly adjusted by 5-7 pixels, and adjustments as large as 15 or more pixels have been necessary (Figure 1).

Data Availability: We are currently generating final mosaics, which will be available in September 2018. Preliminary mosaics and pointing kernels based on the individual networks have been distributed to the planetary science community through the PDS Annex (http://astrogeology.usgs.gov/) in GeoTiff format with available ISIS3 and PDS3 labels, and have been ingested into the JMars software [12].

Future Needs: A globally controlled image data set, and an appropriate base (e.g., altimetry), is necessary for all solid planetary bodies where significant data collection efforts are being made and precision science is desired. This controlled base will allow for the accurate co-registration of all data sets and will enhance the science return and cost-effectiveness of data for that body. For Mars, the THEMIS controlled mosaics, and the associated network, have improved the registration of the THEMIS IR data set and enhanced our knowledge (position, precision, and accuracy) of image placement and small-scale feature location. In addition, these precision products provide for improved targeting for current and future orbital acquisition of data and spacecraft landings and planning of spacecraft surface operations.

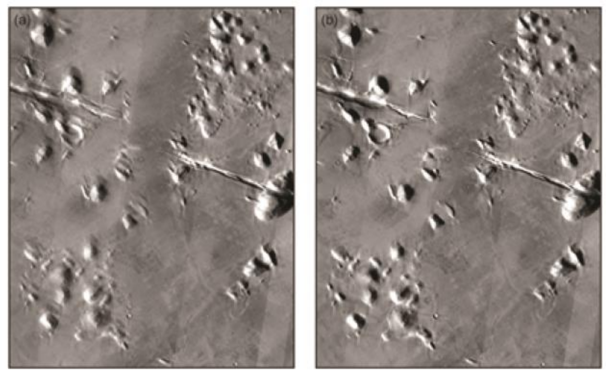


Figure 1. Comparison of uncontrolled (a) and controlled (b) image averaged mosaic products. Portion of the Elysium mosaic, 15.4° N, 162.4° E. A 16-20 pixel shift was necessary to match features in this area. The projection is simple cylindrical with a longitude domain of 0° to 360°.

References: [1] Laura J. R. et al. (2017) *Inter. J. Geo-Info.*, 6, doi:10.3390. [2] Christensen P. R. et al. (2004) *Space Sci. Rev.*, *110*, 85-130. [3] Smith D. et al. (1999) *Science, 284*, 1495-1503. [4] Smith D. E. et al. (2001) *JGR*, *106*, 23,689-23,722. [5] Neumann G. A. et al. (2001) *JGR, 106*, 23,752-23,768. [6] McEwen A. S. et al. (2007) *JGR*, 112, E05S02. [7] Edmundson K. L. et al. (2012) *ISPRS Annals,* I-4, 203-208. [8] Fergason R. L. et al. (2013) *LPS* XLIV, Abstract #1642. [9] Jaumann R. et al. (2007) *Planet and Space. Sci.*, *55*, 928-952. [10] Archinal B. A. et al. (2003) *LPS* XXXIV, Abstract #1485. [11] Archinal B. A. et al. (2004), *XXth ISPRS Congress*. [12] Christensen P. R. et al., *AGU*, Abstract #IN22A-06.

A Virtual Observatory approach to planetary data for Vesta and Ceres. M. Giardino[1] , S. Fonte[2], R. Politi[2], S. Ivanovski[2], A. Longobardo[2], M. T. Capria[2], S. Erard[3] and M. C. De Sanctis[2]
[1]ASI, Rome, Italy (marco.giardino@asi.it - Via del Politecnico, snc - 00133 - Rome, Italy), [2]INAF-IAPS, Rome, Italy, [3]LESIA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, UPMC Univ. Paris 06, Univ. Paris Diderot, Sorbonne Paris Cité

Introduction: We present a data access service in the frame of emerging techniques for virtual observatories.
This service is based upon the IVOA (International Virtual Observatory Alliance) standards widely used among the astronomical community, adapted to the planetary field as defined by the VESPA (Virtual European Solar and Planetary Access) activity in the frame of the European Union-founded Europlanet2020 program [1].
We will discuss the general advantages of such an approach, especially concerning the interoperability and availability of this service.
Some use cases will also be presented to address typical analysis performed by planetary scientists, exploiting the service functionalities and demonstrating the improvement obtained with respect to a more classic approach.

VIR instrument and data format: Data considered here consists of spectral cubes produced by the VIR instrument [2], a visible and infrared spectrometer on board the NASA Dawn mission. The published spectra covers various mission phases targeted to Vesta and Ceres observation in the time frame starting from May 2011 till April 2017: some of these data products are still under revision and will be available for the community at the NASA PDS Small Bodies Node.
The Dawn Mission is a NASA project aiming to identify the mineralogy and the surface composition of the asteroids Vesta and Ceres. The mission payload includes a Visible and InfraRed (VIR) mapping spectrometer combining two data channels in one instrument: the visible and the infrared sensors are housed in the same optical head.
The instantaneous field of view of VIR is represented by a slit of 64 milliradians per 250 microradians, the radiation from the slit is split by the optics to hit both sensors. The products corresponding to a single acquisition are two bi-dimensional frames: for each frame the first axis represents wavelength of the radiation (i.e. the spectral information) while in the other axis signal intensities sampled by the sensor at different locations are stored (i.e. spatial information). The instrument, set to implement a repetition time of acquisition (i.e. temporal information), produces several data frames stored in a 3D structure called cube.
The three dimensions of the cube are referred to as bands, samples and lines, respectively. The bands represent the wavelength of the radiations, the samples are the position on the slit while the lines are the number of acquisitions. A cut along the bands produces a bi-dimensional image of the target observed at the chosen wavelength, while the uni-dimensional array obtained by fixing both sample and line is the spectrum observed in that given point of the target.

Service description: The service is implemented according to the EPN-TAP [4] protocol, which is a restiction of the IVOA's TAP (Table Access Protocol) adapted to Planetary Science. Starting from the original PDS3 dataset of VIR data, we translated this file-based archive structure into a relational database schema representing the same information model, then we ingested each cube metadata into its representation in the relational schema. The used target database schema was derived from a similar data model already defined and used for the VIRTIS instrument on board the Venus Express mission [5]. Then, from the tables obtained in the previous step, the final database view was derived, grouping and formatting the information to match the target schema as defined by the EPN-TAP service specification. Finally, the new service is published on our dedicated TAP server and is declared in the IVOA registry system.

Data Access

Once published in the registry system, this service is easily accessible from any EPN-TAP client, such as the VESPA portal located at http://vespa.obspm.fr, or clients embedded in more general Virtual Observatory applications. The VESPA portal will query simultaneously all EPN-TAP services, favoring cross-examination of data from different services.

In addition, the service can equally be queried by any client supporting the TAP protocol, although with lesser user support level. In this case, the query must be provided in the client interface using the Astronomical Data Query Language, a language based on SQL92. Such clients include TOPCAT, a tool specifically developed for table based data manipulation or Aladin which can plot images and cubes.
In both cases, the resulting data products can be directly displayed in any SAMP (Simple Application

Messaging Protocol)-enabled client. For VIR cubes, a suitable choice is the MATISSE [3] tool, which is web based and can be found at https://tools.asdc.asi.it/matisse.jsp .

Conclusions: The availability of our service will foster the scientific return of the Dawn mission, fulfilling two different achievements: firstly improving the accessibility, usability and interoperability of the data set, secondly promoting the emerging tools and technologies in the virtual observatory domain.

Acknowledgements: The authors gratefully acknowledge the support of the Dawn Instrument, Operations and Science Teams. This work is supported by an Italian Space Agency (ASI) grant and by NASA through the Dawn project.
The Europlanet H2020 Research Infrastructure project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654208.

References:

[1] Erard, S. et al., *VESPA: a community-driven Virtual Observatory in Planetary Science*, Planetary and Space Sciences, in press, ArXiv 1705.09727

[2] De Sanctis M.C., Coradini A., Ammannito E., Filacchione G., Capria M.T., Fonte S., Magni G., Barbis A., Bini A. and Dami M. *The VIR Spectrometer*, Space Science Reviews, Volume 163, Issue 1-4, p. 329-369

[3] Zinzi A., Capria M.T., Palomba E., Giommi P., Antonelli L.A. *MATISSE: A novel tool to access, visualize and analyse data from planetary exploration missions*, Astronomy and Computing, Volume 15, pp 16-28

[4] Erard S., Cecconi B., Le Sidaner P., Berthier J., Henry F., Molinaro M., Giardino M., Bourrel N., André N., Gangloff M., Jacquey C. ad Topf F. *The EPN-TAP protocol for the Planetary Science Virtual Observatory*, Astronomy and Computing, Volume 7, p. 52-61

[5] Politi R., Piccioni G. *SeaBIRD - A VIRTIS-VEX Data repository*, European Planetary Science Congress 2010,held 20-24 September in Rome, Italy, p.401

BROWSING THE PDS IMAGE ARCHIVE WITH THE IMAGING ATLAS AND APACHE SOLR. K. M. Grimes II[1], J. H. Padams[1], A. Stanboli[1], and K. L. Wagstaff[1]. [1]Jet Propulsion Laboratory, California Institute of Technology. {first}.{mi}.{last}@jpl.nasa.gov.

Introduction: The PDS Imaging Node [1] is home to tens of millions of images, nearly 30 million have rich metadata associated with them in the form of PDS labels. Naturally, this metadata varies from image to image, but examples of what it may contain include the following: when it was taken, what spacecraft and instrument took it, and the physical location of the spacecraft when it took the image. This information is stored in the image's label as key-value pairs.

A pressing need of scientists is to locate images with similar characteristics. For some characteristics, such as *mission name*, locating these images by hand may be a trivial process; however, for multi-mission characteristics such as *product type* and *spacecraft clock*, locating images with similar characteristics becomes much more difficult to do by hand. Complicated requests such as "show me all images taken in 2001 with wheels in them" approach impossibility. With the assistance of big data tools, however, such queries may be made.

In order to allow users to search the archive by any available characteristic, the Imaging Node leverages Apache Solr ("Solr") [2], an open-source indexing platform built upon another Apache product, Lucene [3]. The PDS Imaging Atlas ("the Atlas") [4] provides an interactive interface to Solr and allows users to easily and intuitively navigate the archive. Content-based search is enabled via a two-part process: first, a convolutional neural network identifies features in images; next, these features are indexed by Solr so that they can be searched by the Atlas.

Apache Solr: In order to effectively index PDS-compliant images, the Imaging Node creates a *collection* and defines *fields* within a *schema*. Each field in the schema corresponds to a key in the image's label.

Solr provides support for several different types of fields out-of-the-box, including *strings*, *multi-values* (arrays), and *floats*. Custom field types may also be created.

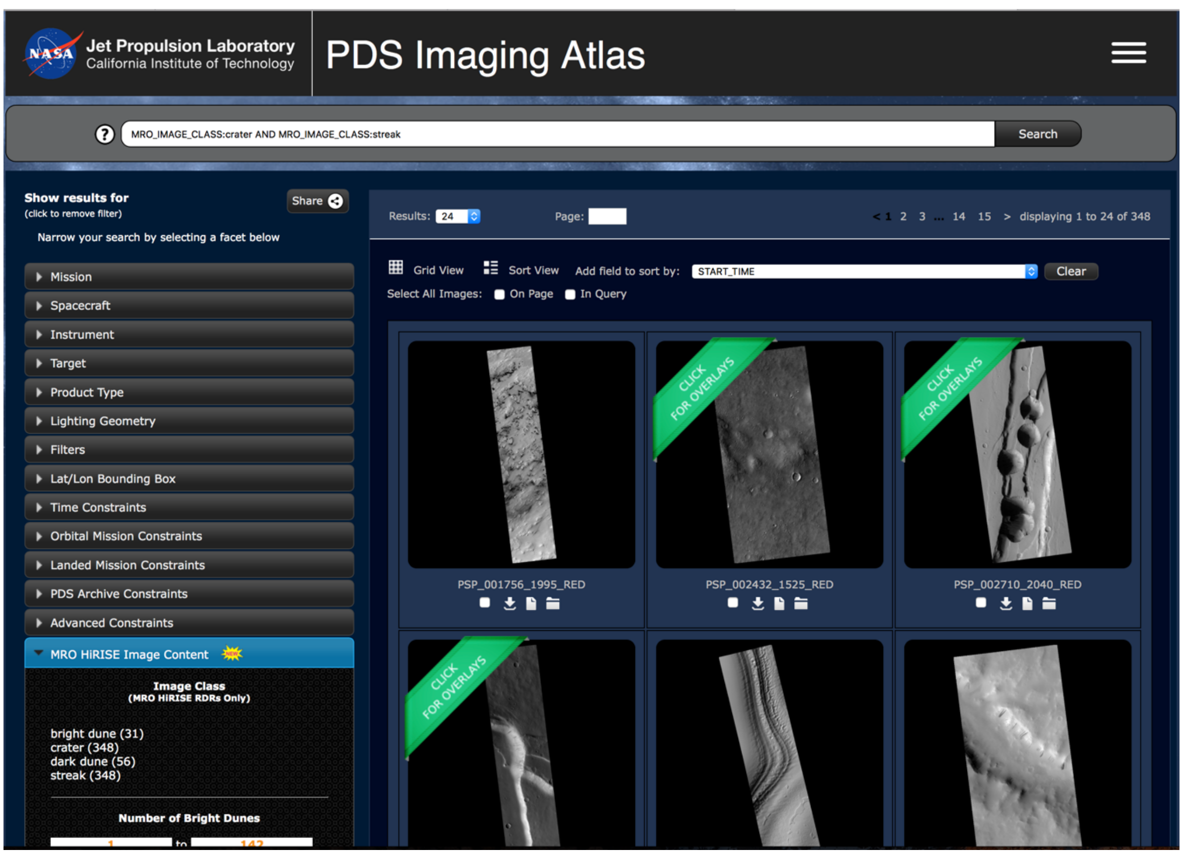


Figure 1: The PDS Imaging Atlas.
https://pds-imaging.jpl.nasa.gov/search

Once the schema has been defined, the main indexing procedure may begin. After some time, the indexing is completed, and the collection is ready to be searched.

Solr Search. One images have been indexed, clients may query for them by any of the fields that they declared in the schema. For example, if a user desires to search for all images that were taken by the Cassini spacecraft and are reduced data records (RDRs), they could make a query to Solr like the following:

```
q=MISSION:CASSINI AND PRODUCT:RDR
```

Query 1: Request for all RDR products produced by the Cassini spacecraft.

Suppose that we want to make a content-based search, for example, of all images taken by the Mars Science Laboratory rover from spacecraft clock counts greater than or equal to 397,000,000 that contain wheels. To do so, we could make the following query to Solr:

```
q=MISSION:"mars science laboratory"
AND SPACECRAFT_CLOCK:[397000000 TO *]
AND MSL_IMAGE_CLASS:wheel
```

Query 2: Request for products produced by the MSL lander after spacecraft clock count 397,000,000 that contain wheels in them.

Solr supports more complicated queries, including nested conditionals and negations.

Browsing the Archive: Because Solr's query syntax can be a bit daunting to users, PDS Imaging Node provides the Atlas as a frontend service to Solr. The Atlas is an interactive interface that allows users to make queries to Solr and interact with the images it returns in an intuitive way.

Facet tabs. Users are able to refine their search by clicking on one of the various facet constraints found in the left hand side view of the Atlas. The available facets to choose from include Mission, Spacecraft, Instrument, and Product Type. When users select a given constraint, a query is made to Solr for documents matching what the user click on.

For example, if the user clicked on the "Cassini" constraint under the "Mission" facet tab and then proceeded to click on the "RDR" constraint under the "Product Type" facet tab, a query similar to *Query 1* above is made to Solr, and matching images are displayed to the user.

Search box. In addition to refining searches using the facet tabs on the left, users may also enter their own custom queries in the search box at the top of the Atlas.

For example, both *Query 1* and *Query 2* could be entered into the search box directly and Solr would return matching documents. This functionality is especially useful in cases where users want to search for constraints on facets that do not have their own facet tab.

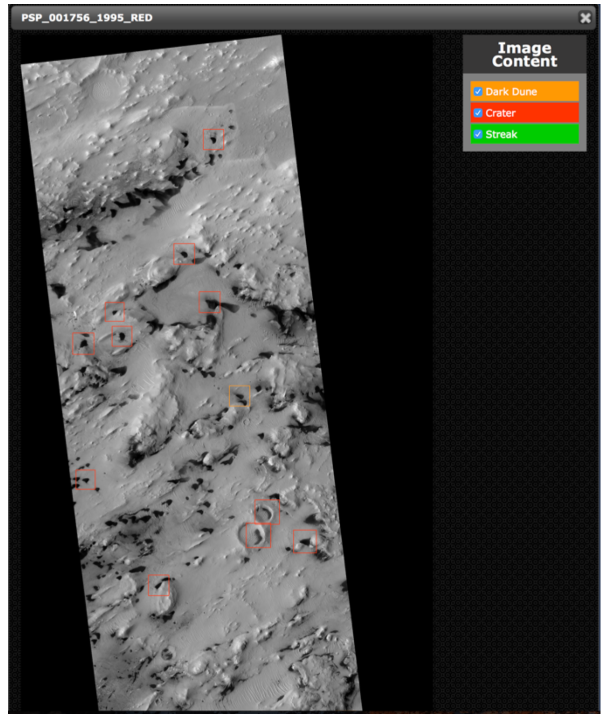


Figure 2: Feature recognition of Martian landscape in the PDS Imaging Atlas.

MRO HiRISE Overlays. In collaboration with machine learning teams at Jet Propulsion Laboratory [5], the Atlas allows users to not only search for images by the features they contain but also to identify these features with bounding boxes on the image.

For example, if a user selects an image taken by the HiRISE instrument that has contains features such as dunes or craters, an overlay will be displayed on top of the image indicating where these dunes and craters are. An illustration of this functionality can be found in *Figure 2*.

References: [1] Planetary Data Systems Imaging Node: https://pds-imaging.jpl.nasa.gov. [2] Apache Solr: https://lucene.apache.org/solr. [3] Apache Lucene: https://lucene.apache.org. [4] PDS Imaging Atlas: https://pds-imaging.jpl.nasa.gov/search. [5] Jet Propulsion Laboratory: https://jpl.nasa.gov.

A Registry for Planetary Data Tools and Services.

S. Hardman[1], M. Cayanan[1], J.S. Hughes[1], R. Joyner[1], D. Crichton[1], E. Law[1], [1]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, CA 91109, Sean.Hardman@jpl.nasa.gov, John.S.Hughes@jpl.nasa.gov, Ronald.Joyner@jpl.nasa.gov, Daniel.J.Crichton@jpl.nasa.gov, Emily.S.Law@jpl.nasa.gov.

Introduction: During the 2015 Planetary Data Workshop, the PDS Engineering Node team presented a plan [1] for taking the prototype Tool Registry developed by the International Planetary Data Alliance (IPDA) and upgrading it by increasing the visibility and enhancing its functionality along with incorporating the registered tools into PDS data search results. This work has been completed with the application deployed into operations by the PDS Engineering Node [2]. Figure 1 below, is a screenshot of the front page of the Tool Registry.

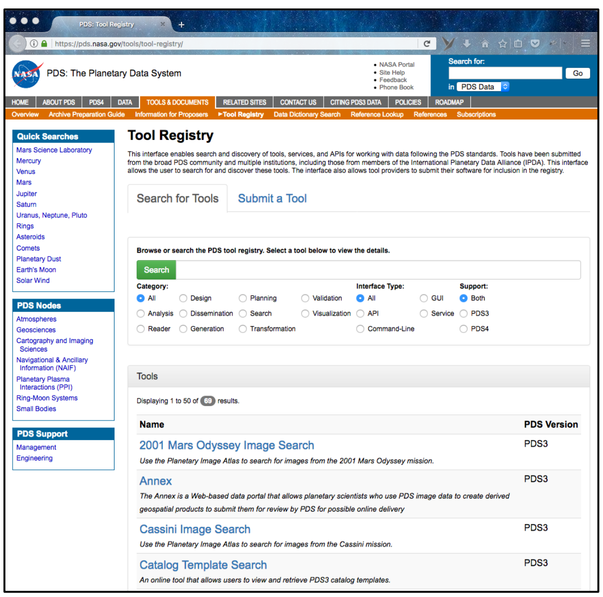


Figure 1: Tool Registry Screenshot

The application enables search and discovery of tools, services, and APIs for working with data following the PDS standards. Tools have been submitted from the broad PDS community and multiple institutions, including those from members of the International Planetary Data Alliance (IPDA). This interface allows the user to search for and discover these tools. The interface also allows tool providers to submit their software for inclusion in the registry.

Along with introducing the planetary data community to the Tool Registry, this presentation will describe and demonstrate how users interact with the application. For those users interested in the details, we will also take a brief dive into the architecture and design behind the application which is built on PDS4 software and the PDS4 information model [3]. The information model provides the framework for capturing metadata that describes the tools and services and the software provides the framework for making this information accessible in the PDS4 data system. Figure 2 below, details the architecture of the PDS4 software utilized to support the Tool Registry.

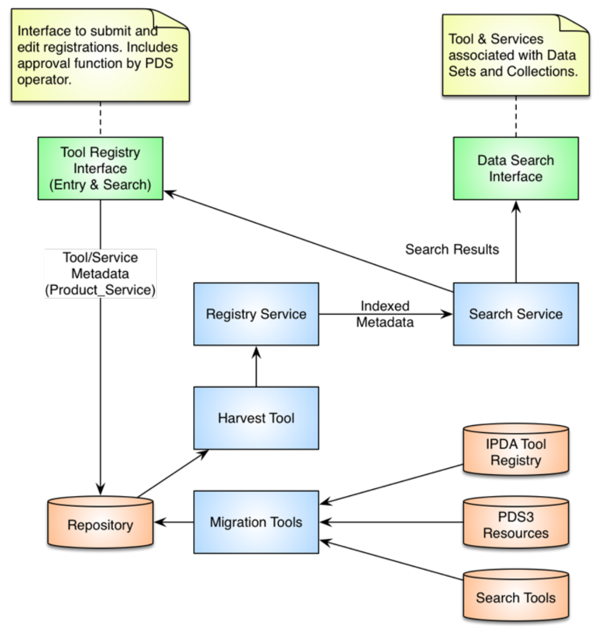


Figure 2: Tool Registry Architecture

References: [1] Hardman S., Hughes J.S., Joyner R., Crichton D., Law E., *Deploying a Planetary Data Tool Registry*, In Proceedings of the 2nd Planetary Data Workshop. Flagstaff, Arizona. June 8-11, 2015. [2] PDS Tool Registry, https://pds.nasa.gov/tools/tool-registry/. [3] PDS4 Information Model https://pds.nasa.gov/pds4/doc/im/current/.

INITIAL PDS4 SUPPORT FOR THE GEOSPATIAL DATA ABSTRACTION LIBRARY (GDAL)

T.M. Hare and L.R. Gaddis, U. S. Geological Survey, Astrogeology Science Center, Flagstaff, AZ, thare@usgs.gov.

Introduction: The NASA Planetary Data System (PDS) has released version 4 of their required archival format [1], now called PDS4. Since 2011, compliance with the PDS4 archiving standard has been required for data archives from NASA-funded missions and smaller research investigations. In contrast to previous versions that used the Object Description Language (ODL) format, PDS4 has been completely re-engineered. The most noticeable change for users will be the requirement for additional metadata and the switch to the eXtensible Markup Language (XML) format. The goals for this move to PDS4 is to improve data discoverability by strongly enforcing consistency in the metadata and allowing for explicit cross-referencing across data products. Herein, we introduce initial support for PDS4 within the Geospatial Data Abstraction Library (GDAL).

GDAL: GDAL, released by the Open Source Geospatial Foundation (OSGeo), offers powerful capabilities for converting and processing geospatial planetary data. GDAL is a format translation library for geospatial raster and vector data [2]. In addition to the newly introduced PDS4 format, GDAL also supports PDS3 (read-only), USGS Astrogeology's image processing formats for the Integrated Software for Imagers and Spectrometers (ISIS2, ISIS3; read/write), Video Image Communication and Retrieval (VICAR; read) format [3], and the Flexible Image Transport System (FITS; read/write) format [4]. Mapping applications, which use the GDAL library for raster I/O (Input/Output), can also directly access these formats. Because GDAL supports more than 100+ formats, it is widely used across the planetary science community to more easily share data. For applications that do not use GDAL for I/O, the bundled routines released with GDAL can be used to convert these formats into more universal geospatial formats (e.g., GeoTIFF).

PDS4 in GDAL Highlights: The NASA PDS Cartography and Imaging Sciences Node (aka "Imaging") has partially funded development of GDAL routines to support I/O reading and data format translations for accepted PDS4 image products (see Section 4.2.1 in the PDS4 Data Providers Handbook, https://pds.nasa.gov/pds4/doc/dph/current/PDS4_DataProvidersHandbook_1.7.0_170517.pdf).

- **Initial Release:** Initial PDS4 support for GDAL was released for testing in September of 2017 (http://www.gdal.org/frmt_pds4.html). Hobu Inc. completed this first of two planned contracts.
- **Templates for PDS4 Creation:** Currently GDAL only supports the translation of the physical properties of the input image (lines, samples, bit type, etc.) and map projection parameters. The PDS4 GDAL driver alone lacks the ability to create or propagate much of the required metadata for writing a PDS4 compliant data product. Thus a complete solution using GDAL will require a well-developed PDS4 XML template or additional scripts will be necessary to retrieve specific values (e.g., from exiting file labels such as those from PDS3 or ISIS3 images), to calculate, translate and otherwise add additional metadata. These PDS4 XML templates must meet PDS4 formatting standards and organization and provide information not propagated by GDAL during format conversion (e.g., author, institute name, processing details). Further, use of PDS4 XML templates must be tailored for each mission and/or image product type. At the current time, the collection of such label templates for a wide variety of image products is minimal but with new or updated examples being regularly added (see: https://pds.nasa.gov/pds4/about/portal.shtml).

  To help support more automated creation of a compliant PDS4 label during a format conversion, GDAL supports user-defined template variables. This allows the user to update these defined variables (e.g., start time, mission phase, etc.) via the command-line or in a scripting language.
- **Remote PDS4 Templates:** To facilitate PDS4 template sharing, GDAL supports loading remote templates from an http address. Thus, PDS nodes, mission teams, and researchers will be able to host example templates from their own website or from sites like Github.
- **Low-level GDAL API:** Full XML access during a format conversion is made available using C++ (xerces library) or a Python (lxml library) application protocol interface (API). This low-level access to the PDS4 label should allow for the development of more robust software applications other than simple scripts.
- **Interoperability:** Because of the enforcement for extensive metadata, it will be difficult for the PDS4 format to be widely interoperable across many applications. To facilitate access to PDS4-

compliant images from a wide variety of application software, GDAL supports writing a detached PDS4 label that references a raw pixel stream from an uncompressed planetary GeoTiff (other formats like FITS or VICAR could also be easily added). Because GeoTiff has broad support across many mapping and scientific applications, this allows for straightforward development of a PDS4 compliant archive and supports use of a more interoperable and universally used scientific image file format at the same time.

- **Direct Mapping Support:** By supporting format drivers in GDAL, the PDS4 reader will eventually find its way into GIS applications like QGIS, UDIG GIS, and Saga GIS. There are plans for support to be added to Esri's ArcGIS Pro and possibly ArcMap GIS.

PDS4 in GDAL Limitations: There are several limitations to the current PDS4 support with GDAL.

- Because GDAL is a geospatial library, the current driver is targeted at supporting high-level derived (map projected) data sets. Low-level (engineering or EDRs) data sets can be supported by using PDS4 XML templates.
- There is currently no PDS4 table support. This will be added during the planned second contract. Like the GDAL PDS3 driver, if a "Latitude" and "Longitude" field are specifically defined, the table will be treated as a geospatial vector point layer, which is suitable for direct display or conversion in many mapping applications.
- Not to be confused with multiple band images which is supported, there is currently no support for writing multiple image arrays in one file (called sub-datasets within the GDAL library). This also will be added during the second contract by using an "append" to an existing PDS4 file. This means that support for PDS4 "composite" headers will need to be handled within the initial master PDS4 template.

In summary, the currently envisioned GDAL workflow for supporting PDS4 image format conversions will rely heavily on user-tailored PDS4 XML label templates and/or user scripts or applications. GDAL, first-and-foremost, is a library to write code against, whether it is for reading, writing or translating across the 100+ supported image formats. Thus an archiving workflow for PDS4 will require up-front archive layout, label design, and metadata input from the data provider. When designing the PDS4 driver for GDAL, the challenge was to create a useful tool, which, most importantly, allows for direct read support (for application I/O and conversion), while also supporting a flexible PDS4 creation solution when partnered with PDS4 XML templates via user scripts and applications.

Challenges for PDS4: Understanding the PDS4 standard and information model remains demanding for users. To keep the format straightforward to maintain within a PDS archive designed for preservation in perpetuity, decisions were made to allow only four basic structural data formats (including 2-D arrays with binary data, tables as repeating records, parsable byte streams and encoded byte streams). Compressed file formats (e.g., Jpeg2000, used by NASA's MRO HiRISE) are no longer allowed. Not only will the PDS4 migrated Jpeg2000 file be significantly larger, beneficial aspects like built-in pyramids (for quick rendering) and streaming capabilities are also not currently supported by PDS4. Because of significant issues like these, PDS3-formatted data may remain available in PDS archives for years into the future.

Support for complicated vector file formats, as those used in GIS mapping applications, are also challenging to support in PDS4. Currently we are researching an XML vector format, the Geographic Markup Language (GML), which is both ASCII-based and yet robust enough to support GIS vector (points, lines, and polygons; see http://bit.ly/2ALQDf0).

Future Updates: While the PDS4 standard has been available for several years, it is still being updated as issues are encountered and capabilities added. We have planned for a future GDAL contract in 2018 or early 2019 to help keep up with these changes and to include capabilities not yet completed. This would include sub-dataset, table and perhaps GML vector-based revisions.

Acknowledgments: This effort was supported by NASA's Planetary Spatial Data Infrastructure (PSDI) InterAgency Agreement and the Planetary Data System (PDS) Cartography and Imaging Sciences Node.

References: [1] PDS4 Information at the Planetary Data System, URL: https://pds.nasa.gov/pds4/. [2] Hare, T.M., et al., 2007, LPSC 39, abs #2536. [3] Gaddis, L.R., Hare, T., and Beyer, R., 2014, Summary and Abstracts of the Planetary Data Workshop, June 2012, U.S. Geological Survey Open-File Report 2014–1056, page 199. [4] Marmo, C. et al., 2018, this volume.

A SANDBOX ENVIRONMENT FOR THE CSM STANDARD AND SPICE.
T. M. Hare and J. R. Laura, U.S. Geological Survey, Astrogeology Science Center, Flagstaff, AZ, 86001 (thare@usgs.gov).

Introduction: Camera or sensor models are a key component of any digital photogrammetric system and are used to accurately project remotely sensed information (e.g., images) to the surface of a planetary body. The U.S. Geological Survey's Astrogeology Science Center (ASC) has begun to develop and test camera models using the Community Sensor Model (CSM) standard [1, 2] in order to support interoperability and broad use of planetary sensor definition across a range of custom and off-the-shelf software tools, as well as a RESTful SPICE web service to remove the need to download, manage, and version spacecraft positional and pointing information.

Herein we present ongoing work ASC is undertaking to provide a programming sandbox environment for employing the CSM standard and associated SPICE information. We define a sandbox as a testing or development environment that allows experimentation outside a given production environment to help demonstrate and teach available capabilities to others.

Background: A sensor model is a mathematical description of the relationship between the three-dimensional object (e.g. a target's surface) and the associated two-dimensional image plane. As described in [3], the quantities needed to define a sensor model can be divided in two broad categories: interior orientation and exterior orientation (or intrinsic and extrinsic matrices from the computer vision literature). The interior parameters are intrinsic to the sensor design and calibration and typically include focal length, location of the principal point, and lens distortions. For more complicated instruments, the interior parameters may also include wavelength dependencies, gain and pixel summing settings, and (for pushbroom sensors) the timing of line exposures and time delay integration (TDI) settings. The exterior parameters describe the location and orientation of the sensor with respect to the target's reference coordinate system. For planetary applications, this information is typically stored in the form of SPICE (Spacecraft, Planetary ephemeris, Instrument, C-Matrix, and Event) kernels and delivered by the Navigation and Ancillary Information Facility (NAIF, [4]).

The CSM Standard: The Community Sensor Model (CSM) Working Group was established by the U.S. defense and intelligence community with the goal of standardizing camera models for various remote sensor types [5]. The CSM standard, now at version 3.0.3, is a framework that provides a well-defined application program interface (API) for multiple types of sensors and has been widely adopted by remote sensing software systems. One of the changes from version 3.0.2 to 3.0.3 was the addition of variable target radii, which enables planetary support (Figure 1). Previously, only an Earth radius (WGS84) was available within the standard.

It is worth noting that the CSM defines a standard interface and does not make the creation of a camera sensor model technically any easier as the implementation details are left to the developer. By defining a standard interface, the CSM supports interoperability between different photogrammetric applications making the development and maintenance of multiple sensor models for similar instruments unnecessary. The CSM API has been designed and continuously tested by terrestrial industry experts and expanded to support necessary planetary parameters. Therefore, we assert that the planetary domain will benefit significantly from the adoption of the CSM standard that has more than decade of design history and development by the CSM Working Group.

Last year our ASC programming team implemented the MESSENGER Mercury Dual Imaging System (MDIS) CSM framing camera model for both the narrow angle and wide angle cameras (NAC and WAC respectively) [2, 6]. In late 2017, we expect to open source and release a pushbroom CSM co-developed by BAE and ASC. This CSM is currently being tested in BAE's SOCET GXP for deriving digital elevation models from MRO's HiRISE and CTX cameras [7] and LROC NAC cameras [8]. This CSM will also be updated to handle HRSC [9].

CSM Sandbox: CSM code is written in C++ for both performance and use of legacy code. In order to facilitate broad use in exploratory environments, improve testing, and support rapid prototyping, we have wrapped the CSM using Cython (http://cython.org/). By wrapping in Cython, we provide Python bindings that can be widely used within the scientific python computing ecosystem (and assert that planetary science community adoption of python continues to increase rapidly [11]). The effort applied to wrapping the CSM has four benefits. First, all tests are written in Python using the PyTest framework that provides support for high level object mocking and easy integration into continuous integration environments. Second, Python bindings all use of the CSM within the near ubiquitous Jupyter notebook [10] environment. The Jupyter Notebook is a backend python kernel and web server with a browser based frontend that allows users to share code, equations, visualizations, and any associated documentation (http://jupyter.org/). Third, using Python provides a single high-level language to access SPICE (SpiceyPy; https://github.com/AndrewAnnex/SpiceyPy) and support for planetary images using the Geospatial Data Access Library (GDAL; http://www.gdal.org). Finally, Jupyter notebooks were heavily used to prototype, in

Python, the algorithms for the C++ implementation of the CSM standard.

CSM Availability: The ASC CSM implementation is available throughout the development process under an open source, public domain license that support maximum reuse by the community. The underlying CSM library (in C++) is built in a continuous integration environment and available via the anaconda (conda) package manager. The ASC maintains a public facing build of the CSM. The Cython wrapper is available via the ASC GitHub website (https://github.com/USGS-Astrogeology/CSM-CyCSM) and via a conda installable package for, at the time of writing, linux-64 and OSX (Windows builds are planned). Finally, the MDIS-NAC and MDIS-WAC implementations are available via both distribution mechanisms (http://bit.ly/CSM-CameraModel). We intentionally separate the underlying library (C++), the python wrapper (Cython) and our implemented camera models (C++/Cython) to support modularity and standard separation of concerns. Usage examples, as we envision a developer or end-user performing exploratory analysis, are available as Jupyter notebooks in our CSM-SET (Sensor Exploitation Tool) repository (http://bit.ly/CSMSET_Jupyter). Currently the MDIS CSM is highlighted but we expect pushbroom CSM examples to be available soon.

The last step to realize our CSM sandbox environment is to more seamlessly be able to access the NAIF supplied SPICE via a RESTful web service. A proof-of-concept prototype has been started called RESTful SPICE (https://github.com/USGS-Astrogeology/Restful-Spice). It currently depends on the Python libraries spiceypy, flask, and numpy. REST is acronym for REpresentational State Transfer and RESTful web services can be thought of as micro-transactions to access simple http addresses. The goal for this implementation is that given an image or stereo-pair, allow the user to simply request the ISD (Image Support Data) which contains the positional description for each image as required by the CSM. With the ISD in hand, the CSM can now be fully tested within the sandbox environment.

Conclusion: Prototype development and subsequent adoption of the CSM standard is the first step in realizing highly interoperable sensor models that can be used and shared across NASA's and international planetary missions. Using the CSM and SPICE web service within an interactive Jupyter Notebook, we find an ideal exploratory environment for sensor model development, data analysis, validation, portability, and finally the capability of demonstrating results to collaborators.

Acknowledgments: This effort has been supported by NASA's Planetary Spatial Data Infrastructure (PSDI) interagency agreement.
References: [1] Hare T. M. and Kirk R.L., 2017, LPSC XLVIII, abs #1111. [2] Hare T. M., et al., 2017, 3rd Planetary Data Workshop, abs #7130. [3] National Geospatial-Intelligence Agency, (2011), Frame Sensor Model Metadata Profile Supporting Precise Geopositioning, NGA.SIG.0002_2.1. [4] Acton, C.H. 1996, Planetary and Space Science, Vol. 44, No. 1, 65-70. [5] Community Sensor Model Working Group, (2010), Community Sensor Model Technical Requirements Document, v. 3.0, NGA.STND.0017_3. [6] Hawkins, S.E., Boldt, J.D., Darlington, E.H. et al. Space Sci. Rev., 2007, 131: 247. doi:10.1007/s11214-007-9266-3. [7] Fergason, R.L. et al., 2016, Space Sci. Rev., 1572-9672, doi:10.1007/s11214-016-0292-x. [8] Burns, K. N., et al., 2012, ISPRS, doi:0.5194/isprsarchives-XXXIX-B4-483-2012. [9] Kirk, R. L., et al., 2017, ISPRS, doi:10.5194/isprs-archives-XLII-3-W1-69-2017 [10] Fernando Pérez, Brian E. Granger, IPython: A System for Interactive Scientific Computing, Computing in Science and Engineering, vol. 9, no. 3, pp. 21-29, May/June 2007, doi:10.1109/MCSE.2007.53. URL: http://ipython.org. [11] Laura, J. R., et al., 2015, LPSC XLVI, abs #2208.

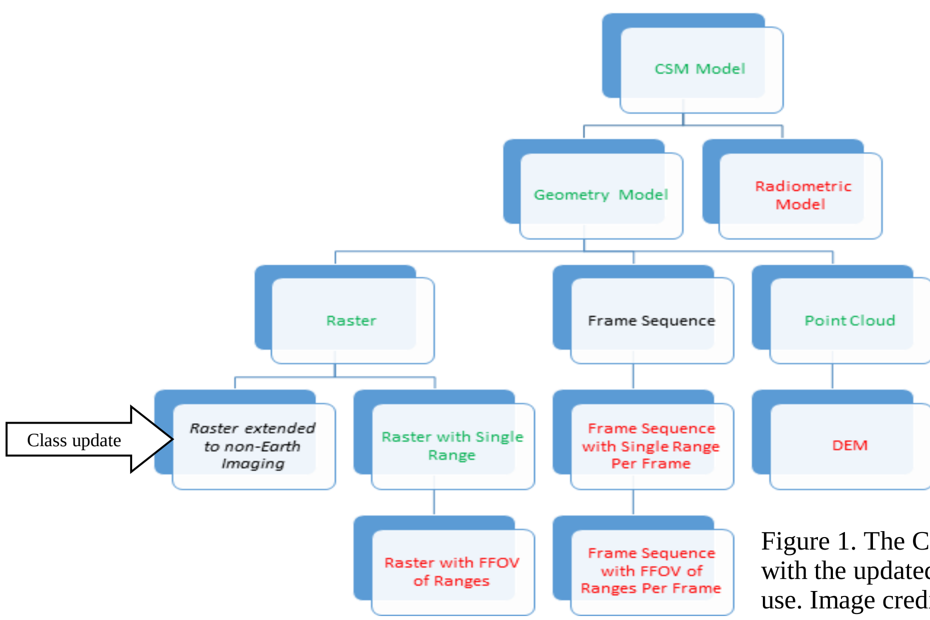


Figure 1. The CSM API hierarchy diagram with the updated class to support planetary use. Image credit: CSM Working Group.

RETRIEVING SINGLE SCATTERING ALBEDOS AND TEMPERATURES FROM CRISM HYPERSPECTRAL DATA USING NEURAL NETWORKS. L. He[1], R. E. Arvidson[2] and J. A. O'Sullivan[1], [1]Preston M. Green Department of Electrical and Systems Engineering, Washington University in St. Louis, St. Louis, MO, USA 63130, [2]Department of Earth and Planetary Sciences, Washington University in St. Louis, St. Louis, MO, USA 63130

Introduction: The Compact Reconnaissance Imaging Spectrometer for Mars (CRISM) on the Mars Reconnaissance Orbiter began acquiring hyperspectral images from 0.362 to 3.92 µm in 2006 [1]. The long wavelength portion of the data have radiative streams that include both solar and thermal terms from the surface and atmosphere. We use the DISORT radiative modeling code to simulate both solar and emission radiative streams, predicting spectral radiance and IOF (radiance/solar radiance) on sensor. We use a neural network (NN) approach to simultaneously retrieve surface spectral single scattering albedos and temperature maps. The NN is trained with numerous laboratory-based spectra chosen to represent the range of possible soils and rocks on Mars, together with use of DISORT outputs that cover a range of SSA and surface temperature values. This approach alleviates the need to assume that any given Martian spectrum is a linear combination of dark and bright area spectra [2][3]. Further, DISORT treats solar surface radiative streams as bidirectional and emission streams as directional hemispherical values for modeling surface SSA spectra, and surface temperatures, gases, and dust and ice opacities. DISORT mapping is implemented as a look-up table with interpolations so it is challenging to retrieve unique temperature and SSA spectra for each pixel.

NN design: *Input and output design.* For NN processing we utilize 491 nodes and for each CRISM scene we input IOF data from 1.4 to 3.85 µm (320 bands), SSA spectra from 1.4 to 2.5 µm (169 bands), which can be uniquely retrieved from the DISORT-based look-up table between IOF and SSA value because temperatures are not relevant in this wavelength region. We aim to estimate single scattering albedo at longer wavelengths and surface temperatures. Therefore 152 nodes with SSA data from 2.5 to 3.9 µm (called SSA_post in Fig. 1), and estimated temperatures are designed as outputs of the NN.

Hidden layers and nodes. According to Funahashi [4] and Hornik [5], any continuous function on a bounded interval can be approximated by a single hidden layer neural network. It is reasonable to assume the inverse of the DISORT look-up table is a continuous function, therefore we use a one layer NN (in Fig. 1). We design the number of hidden nodes as equal to the number of inputs, which is 491. The activation function is chosen as reLU defined as $f(x) = \max(0, x)$ for all hidden nodes, therefore unknown parameters for this NN are the weights on the edges.

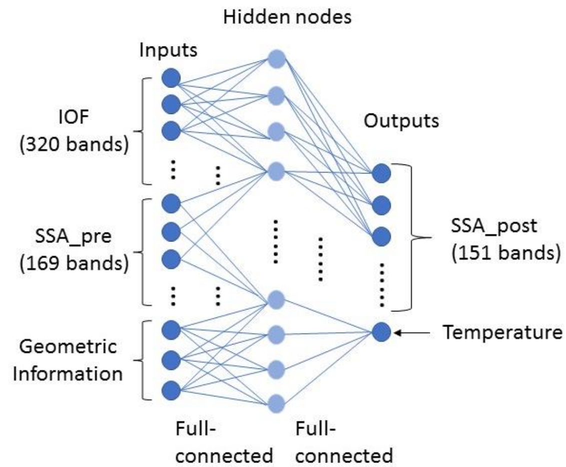


Fig. 1. Flow chart of designed Neural Network.

Dataset. Mars analog laboratory spectral data are used to train the unknown weights in our NN. Because there are~300,000 unknown weights, 300,000 training examples are generated. Each training example contains one lab-based SSA spectrum or combinations of these spectra (320 bands), one temperature, three geometric parameters, and a corresponding IOF cube (320 bands) generated from the DISORT model. SSAs are randomly chosen from the laboratory spectra, temperatures are randomly generated using a reasonable range, and geometric parameters are chosen for each pixel in a scene. We assume that we know the SSA spectral data for 1.4 to 2.5 µm, because, as noted there is a unique mapping from IOF to SSA, i.e., there are no temperature effects at Martian surface temperatures (~230 to 300 K).

Training. To train the NN we estimate the 2.5 to 3.85 µm SSA spectra and temperature for each pixel and compare these to input values, using a backpropagation method (shown in Fig. 2) to minimize the sums of squares of deviations between actual and predicted values. The result is a NN tuned to estimate SSA spectra and temperatures for pixels for each scene. Regularization is needed to avoid NN over-fitting of the training set. An L-2 norm regularization is used and the regularization weight is chosen by cross validation.

Performance analysis: We test our performance for CRISM scene FRS00028346 covering the Curiosity Mars rover landing site and traverse locations.

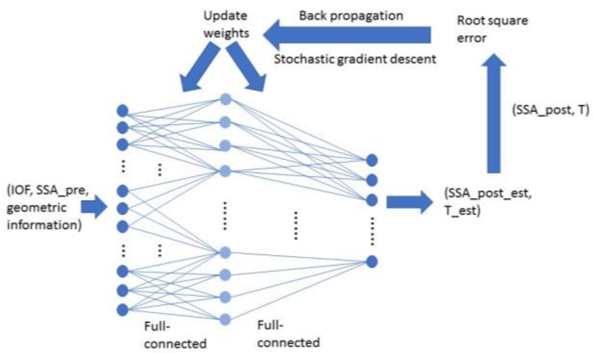


Fig. 2. Backpropagation flow chart.

Predicted Temperature Errors. To explore the sensitivity to temperature predictions we generated a test data set of SSA spectra, trained the NN, and ran the procedure to estimate SSA and temperatures. The average training and test errors are both around 1.3 K for this scene for temperature range 260~285 K. Our NN approximates the inverse of DISORT mapping well.

Real data. The IOF cube for FRS00028346 in Fig. 3 covers Hummocky Plains (brighter area) and part of the Bagnold Dune (darker area) within Gale crater. The temperature mapping from NN for this scene is shown in the bottom of Fig. 4. Another temperature mapping from a thermal model [5] shown in the top of Fig.4 has been calibrated using Curiosity's Remote Environment Monitoring Station (REMS) surface temperature (268 K [6]) of Hummocky Plains in Fig 4. The NN estimate is 268.8 K, which is very close to the measured temperature. Moreover, the NN temperature mapping matches the pattern of the thermal model (Dune Field is warmer than Hummocky Plains). SSA spectra are shown in Fig. 5, retrieved from Dune Field and Hummocky Plains (Fig. 4). The spectra of the Hummocky Plains retrieved from the model and NN-based temperatures are similar for both estimates due to similar temperature estimates. The NN-based spectrum for the Dune Field is different than the one retrieved from the model temperature, likely because the model temperature was estimated at hundred meter scales. The NN estimate is based on 18 m scales and captures more of the sun-facing dune surfaces in this afternoon scene.

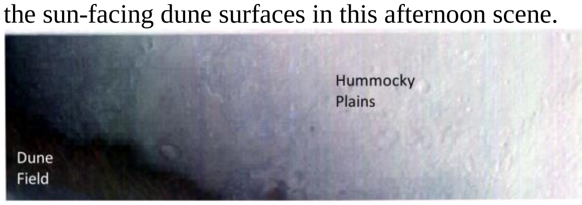


Fig. 3 IOF sensor space cube for FRS00028346 shown in RGB with 1.401 μm, 1.994μm and 2.510 μm.

Future work: 1. We have focused on generating a NN with the corresponding regularization weight to approximate the inverse of DISORT mapping for one scene (fixed atmosphere parameters). In the future, we plan to broaden the scope of the training data to include multiple scenes. 2. We also plan to apply the NN approach to Mars Express Observatoire pour la Mineralogie, l'Eau, les Glaces et l'Activite (OMEGA) [7] data. OMEGA covers ~0.5 to 5.2 μm and the longer wavelengths should provide an excellent way to separate SSA and temperatures. Plans include cross-comparison of CRISM and OMEGA results for the same areas and times.

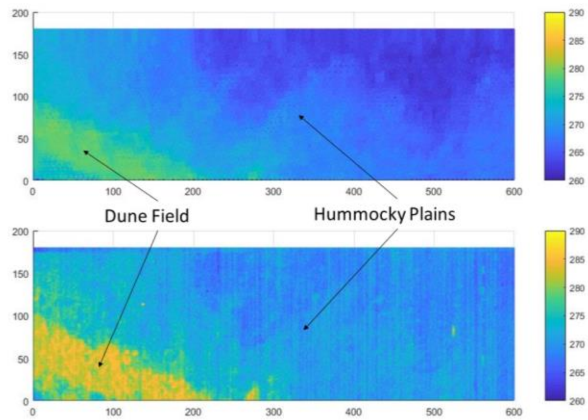


Fig. 4 The top is a sensor space temperature map from the thermal model and the bottom is from the NN. The colorbars in the right show temperature range in Kelvin. Curiosity landed in the area shown at the end of the Hummocky Plains arrow.

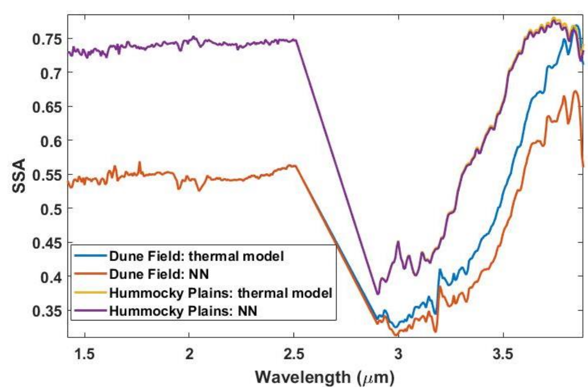


Fig. 5 SSAs retrieved from Dune Field and Hummocky Plains in Fig. 4 for different temperature estimates.

References: [1] Murchie S. et al. (2007) *Journal of geophysical research, vol 112,* Issue E5. [2] Erard S. and Calvin W. M. (1997) *ICARUS, 130,* 449-460. [3] Milliken R. E. et al. (2007) *Journal of geophysical research, vol. 112,* E08S06. [3] Funahashi K. (1989) *Neural Networks, vol. 2,* 183-192. [4] Hornik K. (1989) *Neural Networks, vol. 2,* 359-366. [5] Vasavada A., personal communication, based on albedo and thermal inertia. [6] Sebastian E. (2010) *Sensors 10(10)* 9211-9231. [7] Bibring J. P. et al. (2005) *Science, 307(5717)*, 1576-1581.

APPLYING THE LESSONS LEARNED FROM THE HIRISE GROUND DATA SYSTEM TO THE DEVELOPMENT OF A MODERNIZED GDS FOR CASSIS R. S. Heyd[1], G. A. McArthur[1], R. Leis[1], A. Fennema[1], N. Wolf[1], C. J. Schaller[1], S. Sutton[1], J. Plassmann[1], T. Forrester[1], K. Fine[1], [1]University of Arizona, 102A Sonett Space Sciences Building, 1541 E. University Blvd., Tucson, AZ 85721-0063 (rod@pirl.lpl.arizona.edu).

Introduction: The ground data system (GDS) developed for the High Resolution Imaging Science Experiment (HiRISE) [1], a camera currently flying on board NASA's Mars Reconnaisance Orbiter, is a mature and highly automated data processing system in operation for over 12 years. The system has produced over 5 million image products totaling 160 terabytes of data. The HiRISE technical team is collaborating with the Color and Stereo Surface Imaging System (CaSSIS) [2] instrument team to develop the GDS for the CaSSIS camera on board ESA's ExoMars Trace Gas Orbiter. This provides an opportunity to take many of the lessons learned from HiRISE and apply them to the CaSSIS GDS.

HiRISE Pipeline Processing System: The HiRISE processing system is built on a architecture with multiple computing nodes managed by an application called Conductor developed at the University of Arizona [3]. Conductor operates on "source records" inserted into a database that is polled by an instance of the Conductor application, which launches processing on an individual source. Any number of Conductor instances may be run on a given processing node, with each Conductor responsible for managing a processing task, writing the processing log to a file, and tracking the success or failure of a given processing task. Upon completion of a given processing task, each conductor instance will poll the database for a new source to process and processing task on the next source until the entire list of sources has been exhausted.

The GDS also includes a reporting system indicating processing status and access to completed data products and metadata for the science team.

CaSSIS Processing System: The CaSSIS processing system is currently in development and the tools for managing the GDS have not been fully selected. This paper will compare the requirements of the CaSSIS processing system with the existing HiRISE processing system, examine the lessons learned from the HiRISE and apply them to the case of the CaSSIS which includes different needs and limitations from HiRISE.

References:

[1] McEwen et al. (2007) *JGR, 112,* E05S02, doi: 10.1029/2005JE002605. [2] Thomas, N. (2017) *Space Sci. Rev.*, 212, 3–4, 1897-1944. [3] Castalia, B. (2006) *LPSC XXXVII,* 2159.

THE MULTI-TEMPORAL DATABASE OF PLANETARY IMAGE DATA (MUTED): A TOOL TO STUDY DYNAMIC MARS. T. Heyer[1], H. Hiesinger[1], D. Reiss[1], G. Erkeling[2], H. Bernhardt[1], D. Luesebrink[1] and R. Jaumann[3], [1]Institut für Planetologie, Westfälische Wilhelms-Universität, Wilhelm-Klemm-Str. 10, 48149 Münster, Germany, [2]German National Library of Science and Technology (TIB), Hannover, Germany, [3]German Aerospace Center (DLR), Berlin, Germany. (thomas.heyer@uni-muenster.de)

Introduction: Multi-temporal observations are key to detect and study surface changes and time-critical processes on Mars. Since the 1970s, spacecraft observations have revealed that the martian surface is very dynamic [e.g., 1-3]. The observation of surface changes and processes, including eolian activity [e.g., 4, 5], mass movement [e.g., 6, 7], the growth and retreat of the polar caps [e.g., 8, 9], and crater-forming impacts [e.g., 10] became possible due to the increasing number of repeated image acquisitions of the same surface areas. Today more than one million orbital images of Mars are available [11]. This increasing number highlights the importance of efficient and comprehensive tools for planetary image data management, search, and access.

MUTED is a web-based tool to support the identification of surface changes and processes on Mars. The database enables scientists to quickly identify the spatial and multi-temporal coverage of orbital image data of all major Mars missions. In particular, images can be searched in temporal and spatial relation to other images on a global scale or for a specific region of interest. Additional information, e.g., data aquisiton time, the temporal and spatial context, as well as preview images and raw data download links are available. MUTED is accessible via web at muted.wwu.de and will assist and optimize image data searches to support the analyses and understanding of short-term, long-term and seasonal processes on the surface as well as in the atmosphere of Mars.

Structure: MUTED is based on a three-tier architecture (Fig 1). Metadata of planetary image datasets are integrated from the Planetary Data System (PDS) into a relational database (PostGreSQL) in combination with the PostGIS geospatial extension at the bottom data storage level. In order to provide the multi-temporal coverage, additional information, e.g., the geometry, the number and time span of overlapping images are derived for each image respectively. At the service level, a Geoserver translates the metadata stored in the relational database into web map services (WMS) and web feature services (WFS). WMS provides a global rasterized representation of image coverage. For a region of interest, WFS provides selectable vector representations of the images. Using Common Query Language (CQL), the web services can be filtered by date, solar longitude, spatial resolution, incidence angle, and spatial extend. A GeoWebCache is used to cache map tiles and accelerate as well as optimize the WMS delivery. At presentation level, all services are combined and visualized in the web-based user interface. The user interface was built using HTML, PhP, JavaScript, and Openlayers and provides several features for data selection, filtering, and visualization. A region of interest can be defined based on global spectral, topographic or geologic information. The multi-temporal coverage as well as meta data and the spatial and temporal context of the images are presented on the map, within a timeline or a downloadble feature list.

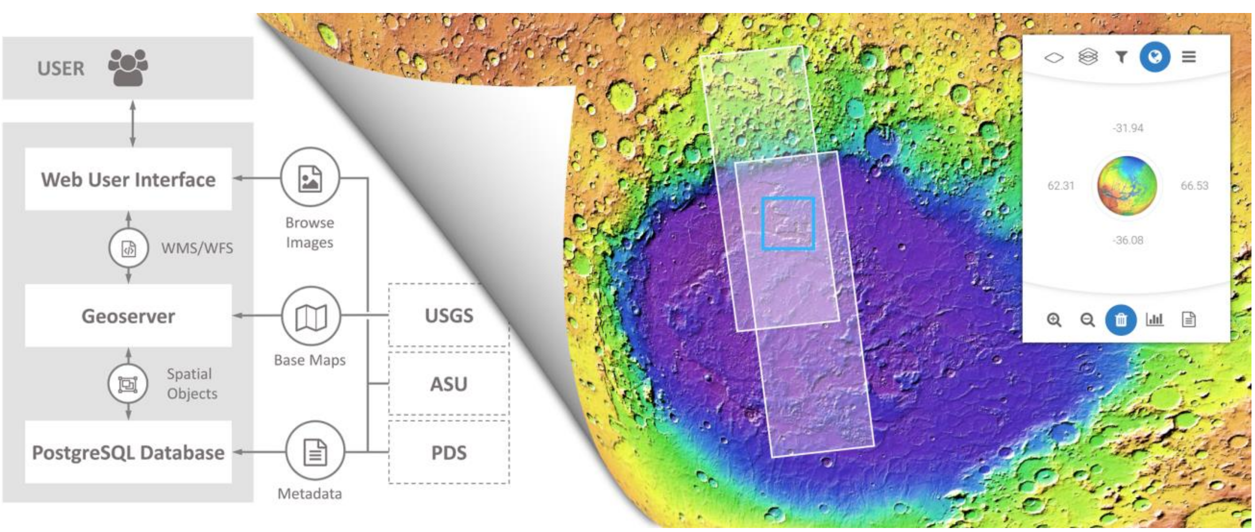


Fig 1.: Architecture of MUTED and corresponding data sources (left) and web-based user interface (right).

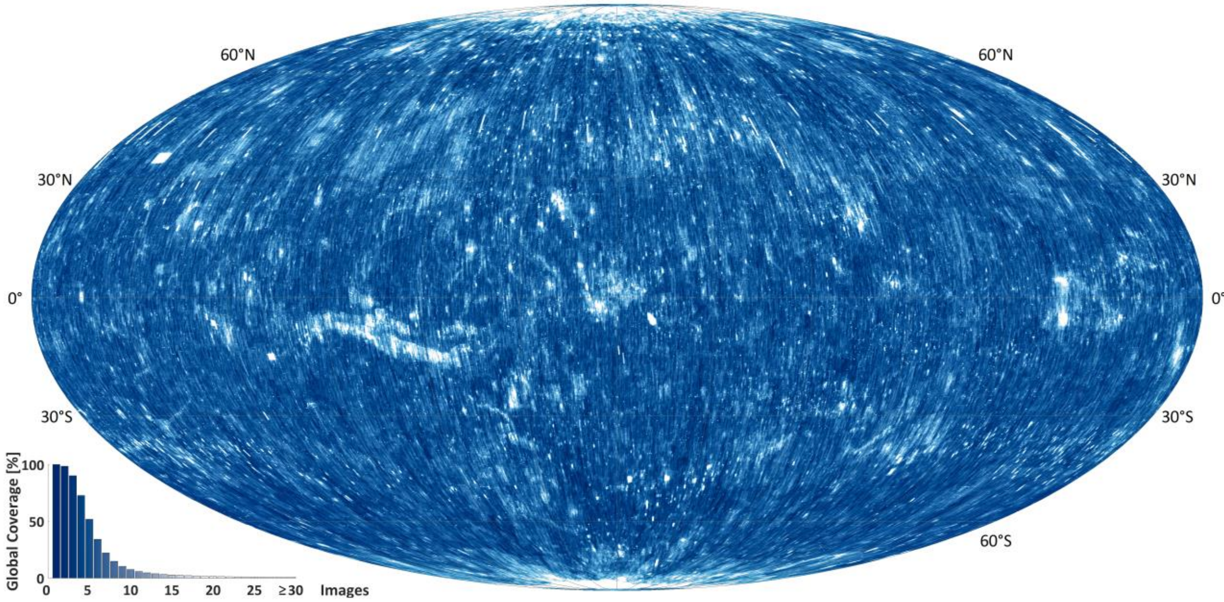


Fig 2.: Multi-temporal coverage of high-resolution orbital images (≤25 m/px) of Mars.

Datasets: At the current state, metadata pertaining to more than 1.27 million orbital images are integrated into the database. The images taken by various instruments including the Viking Orbiter (VO) [12], the Mars Orbiter Camera (MOC) [13] aboard Mars Global Surveyor (MGS), the High Resolution Stereo Camera (HRSC) [14] aboard Mars Express (MEx), the Thermal Emission Imaging System (THEMIS) [15] aboard Mars Odyssey, the Compact Reconnaissance Imaging Spectrometer of Mars (CRISM) [16], the Context Camera (CTX) [17], and the High Resolution Imaging Science Instrument (HiRISE) [18] aboard the Mars Reconnaissance Orbiter (MRO) covering a time range of four decades. The spatial resolution ranges from ~25 centimeters to several kilometers per pixel.

A global coverage analysis reveals that high-resolution images (≤25 m/px) cover 99.9% of the surface of Mars (Fig. 2). Areas with a maximum coverage of ~800 high-resolution images are within the polar regions. Over the last 10 Mars years almost 60,000 high-resolution images were acquired per Mars year with a mean annual coverage of 26.4% of the surface of Mars. While 50% of the surface are covered with at least 5 high-resolution images, the coverage analysis reveals a comprehensive data availability for various change detection tasks. The flexible structure of MUTED allows for a fast integration of upcoming data sets, e.g., from India's Mars Orbiter Mission (MOM) or ESA's ExoMars Trace Gas Orbiter (TGO) mission.

Scientific applications: MUTED enables scientists to explore the multi-temporal coverage of the surface of Mars. The database supports the identification of orbital images and their spatial and temporal context as a basis for various change detection analyses.

In particular, the time span between repeated images can be defined to discover surface changes caused by very short-term and temporally highly variable processes, e.g., dust devils. The difference in solar longitude between repeated images can be specified to observe seasonal changes and processes, e.g., seasonal ice and frost cover. The number of overlapping images can be selected to ensure data availability, e.g., long term changes of the surface of Mars. Due to continuous data acquisition by spacecraft, the amount of image data is steadily increasing and enables further comprehensive analyses of martian surface changes, caused by eolian, mass wasting, polar, as well as impact cratering processes.

References: [1] Sagan et al. (1972) Icarus 17, 346-372. [2] Geissler (2005) JGR 110. [3] Malin et al. (2006) Science 314. [4] Stanzel et al. (2006) GRL 33, L11202. [5] Reiss et al. (2011) Icarus 215, 358-369. [6] McEwen et al. (2011) Science 333, 740-743. [7] Dundas et al. (2015) Icarus 251, 244-263. [8] James et al. (1979) JGR 84, 2889–2922. [9] Calvin et al. (2017) Icarus 292, 144-153. [10] Dauber et al. (2013) Icarus 2013, 506-516. [11] Heyer et al. (2017) LPSC 48, 1019. [12] Carr et al. (1972) Icarus 16, 1, 17-33. [13] Malin et al. (2010) Mars 5, 1-60. [14] Jaumann et al. (2007) PSS 55, 928–952. [15] Christensen et al. (2004) SSR 110, 85–130. [16] Murchie et al. (2007) JGR 112, E05S03. [17] Malin et al. (2007) JGR 112, E05S04. [18] McEwen et al. (2007) JGR 112, E05S02.

MODEL-DRIVEN DEVELOPMENT FOR PDS4 SOFTWARE AND SERVICES. J. Steven Hughes[1], Daniel Crichton[2], Stirling Algermissen[3], Michael Cayanan[4], Ronald Joyner[5], Sean Hardman[6], and Jordan Padams[7], [1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA, steve.hughes@jpl.nasa.gov, [2]daniel.crichton@jpl.nasa.gov, [3]stirling.algermissen@jpl.nasa.gov, [4]michael.cayanan@jpl.nasa.gov, [5]ronald.joyner@jpl.nasa.gov, [6]sean.hardman@jpl.nasa.gov, [7]jordan.padams@jpl.nasa.gov.

Introduction: Software and services that access Planetary Data System (PDS) PDS4 data products need to parse product labels to retrieve, interpret, and process the referenced digital objects. Under PDS4 a driving principle is that the product label provide all of the information necessary for these functions to be performed accurately. However, significantly more information is available in the PDS4 Information Model (IM)[1], the controlling document used to define, create, and syntactically and semantically verify the product labels. This additional information in the IM is made available for use, by both software and services, to configure, promote resiliency, and improve interoperability.

Overview of the PDS4 Information Model: As part of its information architecture, the Planetary Data System (PDS) developed the PDS4 Information Model [1]. This model captures the knowledge about planetary science data at several levels of specificity and allows humans and machines to communicate about the data. As shown in Figure 1, this knowledge has been translated and written to system files in several formats including XML Schema and Schematron, JavaScript Object Notation (JSON), Resource Description Framework (RDF), comma-separated values (CSV) files, XML Metadata Interchange (XMI), and OWL Web Ontology/Description Logic (OWL/DL). The information in these files provide information requirements that augment the system's functional requirements.

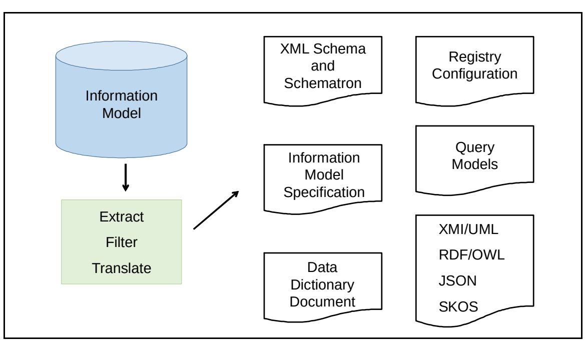


Figure 1 – Generated Artifacts

To address the challenges associated with change within the science community, a multi-level governance scheme was instituted. As Figure 2 shows, the IM has been partitioned into a single common, and several discipline and mission dictionaries.

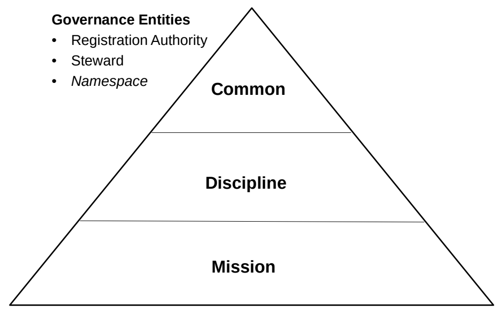


Figure 2 – Multi-level Governance Scheme

The partitioning of the IM reduces the impact of changes, to both the IM and the software and services that use the IM, by localizing the changes to the affected components. Since the common dictionary is designed to be stable, the relatively few changes to this dictionary are localized there. Changes to a discipline dictionary, for example cartography, again only impact that discipline. At the mission level the changes will be much more frequent but again remain localized to a single dictionary. Finally since the model is independent of the implementation, it is insulated from the relatively rapid rate of change in information technology. Software, developed to be configured by or that actively responds to the IM, benefits from this localization with the result that the system, software, and services are more resilient and insolated from changes to the IM. A change in the IM may not even require coding changes.

Multi-level governance also localizes interoperability. For example, the common dictionary enables interoperability across the entire community where each discipline and mission dictionary provides interoperability within a specific community.

Extending the Common Model: The PDS has been successful in developing a consistent and interoperable IM across its diverse disciplines and missions by first providing a stable common dictionary and then providing mechanisms to extend the IM in a controlled supervised enironment. The steward for each of the discipline and mission dictionaries create their individual models by encoding the model into an XML label. The XML label is validated for 'consistency' as part of ingestion into the existing IM (i.e., ensure conformity in the use of data types, units of measurement, and references across dictionaries). Since the initial release of the common model, several discipline and mission models have been released as shown in Figure 3.

Steward Name	Dictionary Description
Common	
Planetary Data System	PDS's common dictionary.
Discipline	
Cartography	Imaging Node's cartography dictionary.
Display	Imaging Node's display dictionary.
Geometry	Geometry dictionary.
Imaging	Imaging node's dictionary.
Ring-Moon Systems	Rings node's dictionary.
Spectral	Spectra dictionary.
Spectral Library	Spectral Node dictionary.
Planetary Plasma Interactions	PPI Node's Wave dictionary.
Mission	
BOPPS	BOPPS dictionary.
InSight	Insight dictionary.
LADEE	LADEE dictionary.
MGS	Mars Global Surveyor dictionary.
MVN	MAVEN dictionary.
OREX	OSIRIS-Rex dictionary.

Figure 3 – Model Dictionaries

Uses Cases: The following use cases illustrate how the exported system files are being used by PDS4 software and services.

Product Label Templates. The chosen implementation for PDS4 product labels is XML. The information model is converted to XML Schema files that contain the classes, attributes, constraints, and relationships defined for each type of data product in the information model. A data provider chooses the appropriate XML Schema files and using an XML editor generates a label template. The data provider then uses the XML editor to populate the template, producing a completed product label. These label templates are also used in product production pipelines to generate a series of product labels for sets of similar products.

PDS4 Validation. The PDS4 Validation tool uses the XML Schema files to validate PDS4 product labels during the ingestion process. This includes validating the order of elements in the label, whether a required element is present, and checking data types and minimum and maximum values. In addition, during generation of the XML Schema files, Schematron files are also written that contain rules to test constraints, for example checking that a value is a member of an enumerated list or that it has been formatted properly.

PLAID. The PDS Label Assistant for Interactive Design (PLAID) tool seeks to simplify and expedite the process of building a PDS4 label template with a simple step by step interface that does not require experience with XML, PDS4 Schemas and Schematron, or knowledge of the specific requirements of a PDS4 product label. PLAID is configured from a JSON formatted file that contains the contents of the PDS4 Information Model. Any change to the model is reflected in the tool.

Conclusion: The PDS4 Information Model, a set of information requirements for science data products, was developed by science experts in the planetary science community. These requirements specify the metadata required to sufficiently describe the data products so that they are scientifically usable by the Planetary Science community both now and into the future. The requirements are also machine readable and can be used to configure software and services that create, validate, and process the data products. Software and services that are written to be configured by or respond to the IM are more resilient, are more readily configurable, and result in reduced maintenance overall. Interoperability is established by the common and existing discipline and mission dictionaries and enhanced over time as new disciplines and mission dictionaries are designed and shared.

References:

[1] Hughes, J.S., Crichton, D., Hardman, S., Law, E., Joyner, R., Ramirez, P., PDS4: A Model-Driven Planetary Science Data Architecture for Long-Term Preservation, IEEE 30th International Conference on Data Engineering (ICDE), Chicago, IL USA, 2014.

Acknowledgements: This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

UPDATE ON THE NASA-USGS PLANETARY SPATIAL DATA INFRASTRUCTURE INTER-AGENCY AGREEMENT.

L. Keszthelyi, J. Hagerty, S. Akins, B. Archinal, M. Bailen, M. Bland, K. Edmundson, R. Fergason, T. Hare, R. Hayward, M. Hunter, J. Laura, S. Sides, M. Velasco. U.S. Geological Survey, Astrogeology Science Center, Flagstaff, AZ (laz@usgs.gov).

Introduction: The phrase *planetary spatial data infrastructure* (PSDI) encompasses a very broad array of activities conducted across the international planetary science community [1]. The data management plans of individual NASA research and analysis grants, the cartography plans of spaceflight missions, the International Astronomical Union, the NASA Planetary Data System (PDS) and its equivalents in other countries, and a host of commercial, academic, and government entities are all key players in this long-term endeavor to provide scientists with the ability to work with data that have a spatial component. This abstract is intended to explain the role of the NASA-USGS PSDI Inter-agency agreement (PSDI-IAA) within this much larger environment. As discussed here, the work funded via the PSDI-IAA is only imperfectly encapsulated within the concept of PSDI.

History: The PSDI-IAA is an evolution of the long-standing NASA Planetary Cartography Program that was housed within NASA's now defunct Planetary Geology and Geophysics Program (PG&G). This program has existed, in one form or another, for well over 30 years. All the functions the USGS carried out under PG&G continue, but new tasks have been added.

Oversight: One key evolution has been the elimination of the external review panel that was the primary oversight mechanism under PG&G. Instead, quarterly and annual reviews are conducted directly between the USGS and NASA. The USGS has developed internal processes to track the progress on each deliverable and assign a status for each task that NASA has directed the USGS to complete. The details of these mechanisms are largely standard project management practices and of little importance to the broader community.

The more interesting problem is the selection of new tasks to take on under this agreement. The expectation has been that the USGS and NASA would make these decisions largely on the basis of community input via the Mapping and Planetary Spatial Infrastructure Team (MAPSIT). MAPSIT ([2] http://www.lpi.usra.edu/mapsit/) is a community advisory group organized to identify strategic PSDI needs for space science and exploration. It is intended to operate in a manner analogous to the NASA Planetary Science Division assessment groups (such as MEPAG, OPAG, SBAG, etc.). While MAPSIT has been slower to stand-up than originally hoped, it still remains the avenue to provide community input on what tasks are high priority items to be completed under the NASA-USGS PSDI-IAA.

Current and Near-term PSDI-IAA Tasks: The inter-agency agreement directs the USGS Astrogeology Science Center to fill certain key needs for NASA planetary exploration, especially as related to PSDI for research and analysis. The work package is determined annually and reviewed quarterly. For FY17 and FY18, the PSDI-IAA has 6 themes: Infrastructure and Data Access (15-16% of funding), Standards (20-21%), Software Development and Tools (35-37%), Products (7-9%), Community Engagement (12-14%), and Management and Personnel (7-9%). The tasks often cross theme boundaries so the level of effort on each type of work is only approximately matched by the dollar amounts in each theme. Note that there currently is not a one-to-one correspondence between the PSDI-IAA themes and SDI themes [1].

Infrastructure and Data Access. This theme includes the following tasks: (1) Database/Dataset Searching Web Services which aims to provide data to a variety of web services via a standard protocol; (2) DPW Management and Development which maintains the hardware and software for the digital photogrammetric workstations that produce digital topographic models of planetary surfaces; (3) MRCTR GIS Lab and Mapping Standards which maintains a guest facility for geologic mappers to work at the USGS in Flagstaff and works to develop and promulgate international GIS standards for planetary science; (4) USGS Regional Planetary Information Facility (RPIF) which is in a slow transition to becoming a network of locations to obtain expert assistance in accessing and working with digital planetary data; (5) New Centralized Map-based Search Portal on Main Astrogeology Website which failed to be completed in FY17; and (6) Astrogeology Geologic Materials Collection starting in FY18 to maintain drill core and other geologic materials from key terrestrial analog sites in a state that is easy for researchers around the globe to access.

Standards. This theme has the following tasks: (1) Planetary Geodesy which coordinates with the IAU Working Group on Cartographic Coordinates and Rotational Elements to provide the community with internationally agreed upon coordinate systems for all planetary bodies; (2) Planetary Geologic Mapping Program Coordination which guides NASA planetary geologic mappers through the process of producing and publishing USGS series map products for planetary bodies; and (3) Planetary Nomenclature which coordinates with the IAU Working Group for Planetary System Nomenclature to provide internationally agreed upon names for features on planetary bodies.

Software Development and Tools. This theme includes the following tasks: (1) Automated Image Matching to Support Control Network Generation which is researching new matching methods in FY17 and is transitioning to application to generating CTX controlled mosaics on a regional scale; (2) Camera Model Architecture which is researching the use of the Community Sensor Model standard to allow planetary data to work with a wide variety of tools developed for Earth remote sensing; (3) Maintenance and Support of the Integrated Software for Imagers and Spectrometers which keeps the infrastructure in place to develop, support, and maintain the widely used ISIS3 package [3]; (4) Integrated Photogrammetric Control Environment (IPCE) which aims to bring a major improvement in the usability of the tools used to apply rigorous photogrammetric control to products from the full variety of planetary data [4]; (5) SOCET GXP Conversion which is migrating planetary stereogrammetry from the deprecated SOCET SET software to current SOCET GXP software. In FY18 we plan to add the following tasks: (1) Git for Integrated Software for Imagers and Spectrometers to bring the ISIS3 software repository up to USGS required standards and (2) Improve ISIS3 Control Networks to Handle Larger Data Sets which is necessary for products we expect to produce in FY19 and beyond.

Products. This theme includes the following tasks: (1) Completion of the USGS' Enceladus Cartographic Package which adds key supplementary information to the recently completed controlled image mosaic for Enceladus [5]; (2) MESSENGER Image and Topographic Maps of Mercury which will produce for Mercury the type of product that has proven to be very popular with researchers and the public for the Moon; (3) THEMIS Controlled Mosaics of Mars which will complete a multi-year effort in FY18; (4) Creation of a Global Shape Model for Enceladus which is a logical extension of the controlled image mosaic; and (5) Updated Global Basemap and Renewed Photogrammetric Results for Europa to provide the best basis for planning the future Europa Clipper and Europa Lander missions.

Community Engagement. This theme has the following tasks: (1) USGS Review Panel Attendance which covers the full-cost accounting requirements for USGS staff to serve on NASA review and advisory panels; (2) USGS Planetary Spatial Data Infrastructure Community Engagement which includes the administration of MAPSIT activities; and (3) ISIS3 Workshops to restart in FY18.

Management and Personnel. This theme includes the following tasks: (1) NASA-USGS PSDI IAA Implementation which is the internal management of the cornucopia of PSDI-IAA work; (2) PSDI Human Capital Maintenance which provides funding for postdocs and graduate students to obtain real-world experience working on PSDI activities; and (3) Software Committee to incrementally improve the USGS software development processes. Tasks (1) and (3) will be funded by the USGS but the progress on their deliverables will be monitored by NASA. Task (2) has focused on Ohio State University and Northern Arizona University but can be expanded to include other institutions.

Future Directions: By the time of the meeting we should be able to report on the FY19 plans. At this time, the expectation is that the relative effort towards generating products will increase at the expense of software development. This will happen naturally as the SOCET GXP and IPCE activities transition from software development to the use of the new tools to create cartographic products. Longer-term, we are looking to better support data collected from terrestrial analogs and active spaceflight missions.

References: [1] Laura, J.R., et al., (2017) ISPRS Int. J. GeoInf. [2] Lawrence, S., et al. (2016) *LPI* XLVII, Abstract #1710. [3] Sides, S.C., (2017) LPSC 48, Abstract #2739. [4] Edmundson, K.L., et al. (2015) LPSC 46, Abstract #1454. [5] Bland, M.T., et al. (2016) LPSC 48, Abstract #2342.

PROVIDING OBSERVATION CONTEXT VIA KERNEL VISUALIZATION AND INFORMATICS FOR PLANNING AND DATA ANALYSIS. John N. Kidd Jr.[1], Sanford Selznick[2], and Carl W. Hergenrother[3], [1]Ascending Node Technologies, LLC. 2548 E. 4th St. Tucson, AZ 85716, johnkiddjr@gmail.com, [2]Ascending Node Technologies, LLC. 2548 E. 4th St. Tucson, AZ 85716, sanford-psdi@selznick.com, [3]Ascending Node Technologies, LLC. 2548 E. 4th St. Tucson, AZ 85716, carlhergenrother@gmail.com.

Introduction: To fully understand the impact of planetary data, the lighting and viewing geometry of said data must be well known for planning and analysis. Typically this information can be gleaned from SPICE kernel files detailed by NASA's Navigation and Ancillary Information Facility (NAIF) [1]. The SPICE data system consists of specific kernel types which describe objects (both planetary and spacecraft) position and orientation as a function of time with respect to the J2000 coordinate system [2]. Additionally, SPICE kernels contain (1) geometric definitions of custom coordinate frames that may be of particular use for a given mission, (2) position, (3) orientation, and (4) field-of-view details for each of the instrument payloads onboard a spacecraft. From these sources, along with custom software, extensive details of an observation (i.e.: viewing angle, lighting angle, range to surface, phase angle, etc.) can be ascertained [2]. For this reason, SPICE files are often the underlying foundation of a planetary exploration mission [2][3]. Figure 1 provides a high level example of the relationships defined within SPICE kernels [1][2].

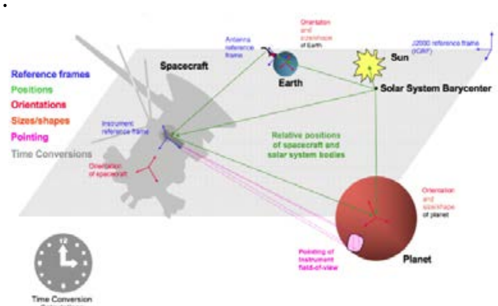


Figure 1: Solar System Geometry as recorded by SPICE kernels [1]

However, the utilization of SPICE kernels requires a not insignificant amount of prior knowledge to use effectively. Generally, planetary exploration missions have a small handful of SPICE experts whom are tasked with assisting other team members' use and understanding of kernels in addition to their regular duties. Additionally, the act of visualizing, verifying, and/or debugging data contained within kernels regularly involves using a gamut of tools and processes to glean partial facts regarding the kernels.

Features: Herein we introduce a novel software system to allow engineers and scientists to rapidly leverage our SPICE expertise and experience through an easy-to-use and configurable application with advanced visualization capabilities. The intent is to make the SPICE architecture more readily accessible to a larger audience of engineers and scientists working on any given mission. To make SPICE as accessible as possible, the proposed tool is designed for online deployment for all users of all platforms. To this end, anyone either on an active mission or using historic data from the Planetary Data System (PDS) can quickly and easily understand the context of a mission's SPICE kernels.

This tool provides two central features: (1) SPICE Visualization and (2) SPICE Informatics and Kernel Debugging.

SPICE Visualization. A common task for many engineers and scientists on a mission is to visualize data contained within kernels to confirm that either a plan that is in work is designed to meet mission goals, or to determine whether or not an observation which has already occurred was compliant with said plan. The use of rapid visualization can allow for more agile operations design by the fact that any team member can quickly verify the validity of an observation design as well as offer more informed and constructive feedback if the design is invalid. An example of such a visualization is provided in Figure 2.

A few COTS tools have attempted to provide this capability to varying degrees of success, but no publically available tool has yet to reach the point where a user can simply upload their kernel pool and immediately generate meaningful visualizations that depict the kernel's content [4][5][6]. By focusing on this specific need, direct kernel visualization is more accessible than it has ever been. Using this tool allows mission engineers and scientists to relax demands upon the team's group of SPICE experts.

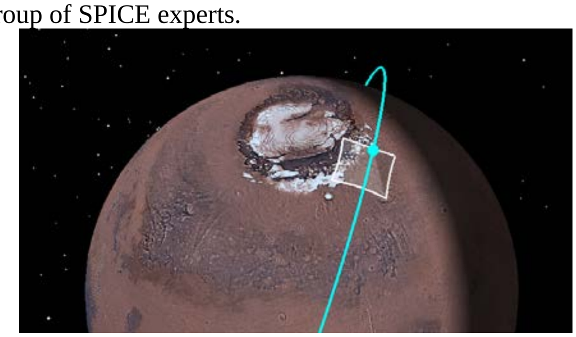


Figure 2: Example visualization of a spacecraft orbiting Mars.

SPICE Informatics and Kernel Debugging. Frequently on a planetary exploration mission, questions come up regarding which kernels contain which data, which kernels link together, and which data sources take priority, and at what time, depending upon the coverage intervals. Sometimes these questions have simple answers, but more often than not a much more involved answer is necessary. To address these issues, scientists and engineers draft either a long email or a rough sketch of what the kernels contain.

The proposed tool offers a more thorough and robust response: interactive visualization of the kernel content combined with relationships to other kernels can be computed on the fly from a given meta-kernel, as demonstrated in Figure 3 and Figure 4.

A common pitfall experienced while using the SPICE architecture is that a particular coordinate frame or reference data is not available within a meta-kernel to allow a query to be computed. In response to this, the software is designed to provide insight regarding a failed connection in a query. An example of this is depicted in Figure 3 below in which the state of a spacecraft object, "MRO" was queried with respect to the "SN_FRAME" coordinate frame at planet "Mars". However, a SPK doesn't exist providing Mars' J2000 state with respect to the Sun, and thus the query fails.

Figure 4 provides timelines showing which kernels are informing a spacecraft's trajectory and attitude as a function of time. Further extensions of this capability can include additional information becoming available in scroll-over windows.

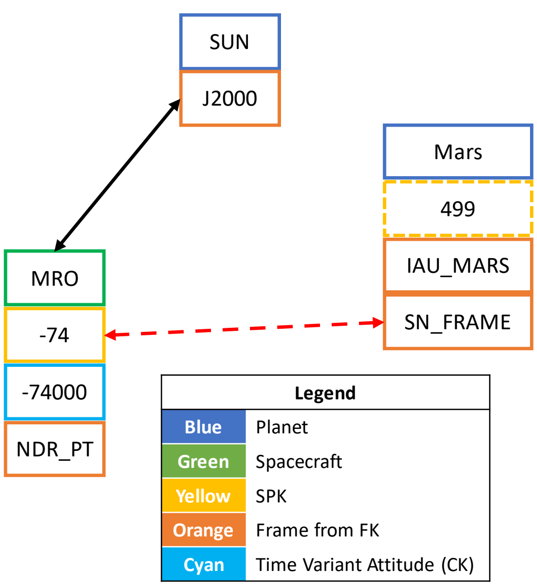


Figure 3: Example SPICE kernel flow diagram.

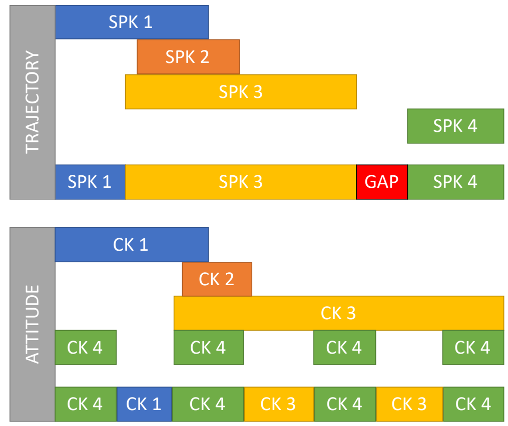


Figure 4: Example SPICE kernel data source timeline.

Summary: From our lessons learned and SPICE expertise we have laid out the features and capabilities of a new web-based tool to provide an accessible platform to obtain context and informatics from a planetary mission's SPICE kernels.

With modern, agile development practices, the foundation is laid to be easily extendable to service the particular needs of any mission. This will be greatly beneficial to a mission operations team, science team, and especially the efficiency of the interaction between the two.

Additionally, the proposed tool provides key functionality to scientists whom are analyzing data products already delivered to the PDS. In this circumstance, individuals familiar with the mission and the details of its kernels may no longer be readily accessible; this tool can instead provide these insights.

References:

[1] "Ancillary data services of NASA's Navigation and Ancillary Information Facility", Acton, Charles H., Planetary and Space Science, Volume 44, Issue 1, p.65-70. Jan 1996. [2] "SPICE Products Available to the Planetary Science Community", Acton, C.H., 30th Annual Lunar and Planetary Science Conference, March 15-29, 1999, Houston, TX. [3] "What's New in SPICE", Liukis, M., Third Planetary Data Workshop and The Planetary Geologic Mappers Annual Meeting, June 12-15, 2017, Flagstaff AZ. [4] https://naif.jpl.nasa.gov/naif/cosmographia.html [5] "WebGeocalc and Cosmographia: Modern Tools to Access SPICE Archives", Semenov, B. V.; Acton, C. H.; Bachman, N. J.; Ferguson, E. W.; Rose, M. E.; Wright, E. D., Third Planetary Data Workshop and The Planetary Geologic Mappers Annual Meeting, June 12-15, 2017, Flagstaff, AZ. [6] "How STK Automatically Loads SPICE Files", http://help.agi.com/stk/index.htm#stk/SPICE.htm?Highlight=spice

DEMONSTRATION OF UPDATED OLAF CAPABILITIES AND TECHNOLOGIES. C. Kingston[1], E. Palmer, J. Stone, M. Drum, C. Neese, B. Mueller, [1]Planetary Science Institute, 1700 E Ft Lowell Rd, 110, Tucson, AZ, ckingston@psi.edu.

Introduction: The On-Line Archiving Facility (OLAF) provides an interface through which users can submit data to the NASA Planetary Data System (PDS) so that it can be reviewed, archived, and made available to the public. OLAF is maintained by the PDS Small Bodies Node.

Goals. The PDS4 standard can be complex and intimidating to learn, making archiving data a burdensome task. OLAF is designed to make submitting data to the PDS easier. OLAF outputs properly-formatted and PDS4-compliant XML product labels for several data types (with more currently under development). However, users are not required to maintain current understanding of an evolving standard and are thus isolated from PDS4 terminology altogether.

In support of these goals, OLAF has been undergoing a transformation into a more modern web application. The most recent improvements to OLAF include:

- developing a separate client application using Angular.js and Node.js
- separating client development from server logic
- developing a RESTful API from the existing OLAF codebase to consume data supplied by the server
- providing tools to manage observing systems and instruments/telescope relationships without having to learn terms specific to PDS4 or OLAF

OLAF includes the capability to upload tabular data as Comma Separated Values (CSV) files, as well as an improved method for uploading and batch processing to simplify the data submission process.

Benefits: These are some of the benefits of using the Angular framework with OLAF:

- Simplifies and enhances client-side development
- Establishes an extensive library of HTML modules with associated functionality
- Promotes re-usability of code blocks and custom validations
- Minimizes existing code revision
- Uses highly-readable JSON strings for data exchange
- Improves code maintainability
- Provides access to hundreds of thousands of open-source software packages
- Incorporates new data types easily

CSV Tables. The use of CSV files for tabular data allows users to use spreadsheet software like Excel to generate data files. CSV files also make it simpler to embed header definitions directly in the files either by hand or programmatically. Including the metadata in this way significantly reduces the number of steps in the data upload process. CSV files are also easier to format and are more human-readable than fixed-width tables.

A GIS Layer Depicting Proposed Human Landing Sites & Exploration Zones on Mars & Tool to Investigate These. Kara Latorella[1] and Matthew Tisdale[2], Lavontria Aaron[3], John Leinenveber[3], Joseph Wilson[3], Alyssa Werynski[3], Phyllindia Gant[3]. [1]Space Mission Analysis Branch, NASA Langley (K.A.Latorella@NASA.GOV), [2]Booz-Allen Hamilton, STARSS III contract NASA Langley,). [3] National Institute for Aerospace, NASA Intern.

Introduction: NASA and associated partners have been selecting landing/impact sites since the 1950s. Typically, this selection occurs as a result of accomplished scientists, engineers, and mission analysts from the international community who submit proposals, analyze these with respect to NASA mission objectives, and attend successive workshops wherein these proposals are discussed and rated. Good candidates meet engineering constraints and enable science/resource mission objectives. Golombek et al. [1] document this selection process for the Mars Science Laboratory. More recently, site selection has begun for the next Mars rover – the Mars 2020 Rover [2].

In October 2015, NASA held the first Human Landing Site Selection (HLS2) for Mars workshop. Forty-two presentations were provided. These were discussed in light of NASA's mission constraints and objectives for science and planetary resource acquisition/utiliza~~i~~tion. This workshop invited participants to "begin the conversation about what constitutes a good landing site for future planetary scientists and astrobiologists" [3] by proposing Exploration Zones (EZs). Each EZ contains a Landing Site (LS), a Habitation Site (HS), and Regions of Interest (ROIs) that address science and engineering objectives – all within a 100km traverse from the centralized LS [4] (Figure 1, [5]). Participants characterized EZs' scientific and

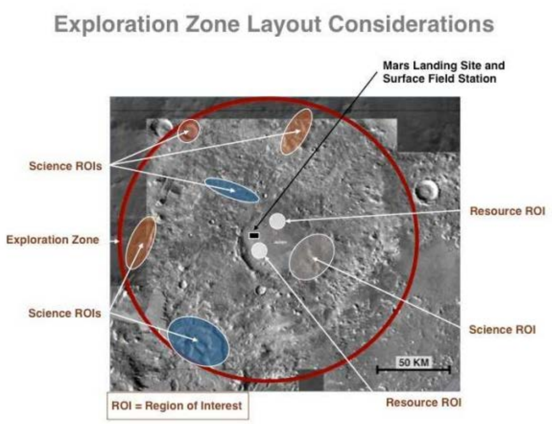


Figure 1. HLS2 Workshop Exploration Zone Example.

engineering benefits in terms of a "rubric" [6] and cited referenced datasets justifying claims. The results of this initial workshop were presentations, video of authors' presentations, and short papers (3 pages). The next steps identified in this workshop were to further discuss the data and analyses required to improve proposed EZs, and consider how the community would best be engaged to continue this conversation [7].

An Integrated GIS Layer & Analysis Environment: The proposals submitted to HLS2 represented a breadth of knowledge from the international scientific and engineering community, and as such contributed an invaluable start to this important conversation. That each proposal was described in separate documents presented an unwieldy dataset for further analysis. HLS2 proposals also contained wide variability that made their data difficult to coalesce. Proposers used different data sources for the same scientific/engineering inferences; used different measurement units, and scales; presented images on different base maps; used different map projections; some used geographical information systems (GIS) to coordinate referenced data, while others simply drew ellipses on base maps, and still others did not include graphical depictions of their proposed site(s). These challenges impede not only NASA's ability to compare and assess contributions, but hinder what could be a valuable collaborative discussion among participants, and the broader scientific/engineering community.

NASA Langley, as part of an internal System Analysis & Concepts Directorate (SACD) and sponsored by the Evolvable Mars Campaign Surface Operations Task, undertook the task of coalescing proposed EZs into a single GIS map layer. The LaRC_Mars_HLS2 Layer depicts all proposed EZs and their constituent LS, HS and ROIs. Each EZ ROI was also associated with rubric characterization data as well as authors' documentation. This effort received generous collaboration from partners in the United States Geological Service (Trent Hare), NASA Ames Research Center (Brian Day), the Jet Propulsion Laboratory (Emily Law) and Arizona State University (Jonathon Hill), who ensured that this effort was compatible with use of data products as produced from others, e.g., USGS data layers (Hare); and other graphical information and viewing software, i.e., MarsTrek (Day & Law); and JMARS (Hill). LaRC_Mars_HLS2 Layer, is available as a shape file that can be imported into standard GIS environments, permitting the assessment of all EZs with a common representation and in consideration of the same contextual data (e.g., geological features, orbital data depicting science-related phenomena). This content will be vetted by proposal authors.

While the LaRC_Mars_HLS2 layer can be used in standard GIS systems, this effort also developed a web-accessible environment, LaRC_MarsGIS_Viewer & Collaboration Tool (LaRC_MarsGIS), based on ESRI Inc. ArcGIS software to support vetting of the layer in a common environment. LaRC_MarsGIS supports display of the LaRC_Mars_HLS2 Layer and permits concurrent display of other Mars data sources, enables map functions (e.g., zoom, pan, measure distance), editing of EZ location and constituent data (by authors), and supports collaboration. A layer file is provided that unifies the depiction and this associated data. The LaRC_MarsGIS also contains ancillary user help documentation and videos. Figures 2 -4 show screenshots from the viewer. The LaRC_MarsGIS will be available for the general public following review of the layer by authors.

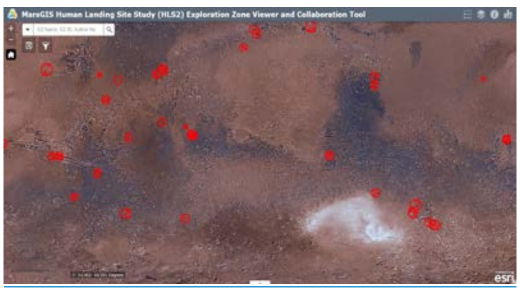


Figure 2. LaRC_Mars_HSL2 layer Overview~~.~~.

Figure 3. LaRC_Mars_HLS2 EZ information.

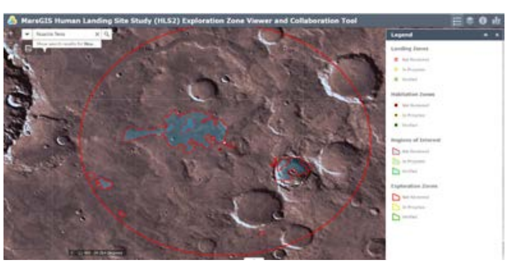


Figure 4. LaRC_Mars_HLS2 author vetting.

Discussion: It was a specific aim of this effort to provide a common representation and common platform to encourage community deliberation for the important task of considering the first human Mars landing site. Group annotation tools support the externalization of judgments and assumptions about information in context, [8] and information tagging facilitate accurate referencing in a common context – especially important for situating group processes that may be asynchronous. Extensions to LaRC_MarsGIS are envisioned to include additional features support collaborative spatial decision support in mission contexts (see [9]).

References:

[1] Golombek, M., et al. (2012). Selection of the Mars science laboratory landing site. *Space Science Reviews*, *170*(1-4), 641-737.

[2] NASA (2017). *Scientists Shortlist Three Landing Sites for Mars 2020*. (accessed 11/3/2017). https://www.nasa.gov/feature/jpl/scientists-shortlist-three-landing-sites-for-mars-2020.

[3] Grunsfeld,J.M. (2015) *HLS2 October 2015 workshop – Mars Human Landing Site Workshop.* (accessed 11/3/2017). https://www.nasa.gov/sites /default/files/atoms/files/grunsfeld-151027-mars-human-landing-site-workshop-final-jmg_tagged.pdf.

[4] NASA (2015a). *HLS2 October 2015 workshop – Supplemental Background Information*. (accessed 11/3/2017).https://www.hou.usra.edu/meetings/explorationzone2015/program_presenter_info/Supplemental%20_Paper.pdf (accessed 11/3/2017).

[5] NASA (2015b). *HLS2 October 2015 workshop – Announcement.* (accessed 11/3/2017). https://www.nasa.gov/sites/default/files/atoms/files/hls2-2nd-announcement-7-8.pdf (accessed 11/3/2017).

[6] NASA (2015c) *HLS2 October 2015 workshop - Exploration Zone Rubric*. (accessed 11/3/2017). https://docs.google.com/spreadsheets/d/1WMBpr7btX4-WZ2u0dhTdEQz1oVsnkLRk7oqKqhmJ3k/edit#gid=0

[7] Bussey, B., & Davis, R. (2015) *HLS2 October 2015 workshop - Human Landing Sites Study (HLS2) Group Discussion*. (accessed 11/3/2017). https://www.nasa.gov/sites/default/files/atoms/files/hls2-group-discussion-fridaytagged.pdf

[8] Convertino, G., et al. (2008). Articulating common ground in cooperative work: content and process. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1637–1646..

[9] Sugumaran, R. & DeGroote, J. (2011). *Spatial Decision Support Systems.* CRC Press: NY,NY.

THE RELATIONSHIP BETWEEN PLANETARY SPATIAL DATA INFRASTRUCTURE AND THE PLANETARY DATA SYSTEM. J. Laura[1], R. E. Arvidson[2], and L. R. Gaddis[1] [1]USGS Astrogeology Science Center, Flagstaff, AZ, [2] Dept. Earth & Planetary Sciences, Washington University, St. Louis, MO (jlaura@usgs.gov).

Introduction: The recent codification of a theoretical Planetary Spatial Data Infrastructure (PSDI) framework [1] and concurrent presentation of PSDI themes at several planetary science conferences [2, 3, 4] and advisory group meetings [5, 6, 7] have engendered two consistent questions: (1) how is PSDI different from the Planetary Data System (PDS) and (2) what is the relationship between PSDI and the PDS? This abstract addresses these questions from the perspective of PDS science discipline node lead scientists and PSDI experts. While the PDS provides a valuable planetary data archive service and has taken steps towards implementing aspects of PSDI, future PSDI development will play a significant role in addressing user needs that then propagate into archival systems such as the PDS.

Planetary Spatial Data Infrastructure: A PSDI is the collection of users, policies, standards, data access mechanisms, and the data proper [8, 9]. PSDI components are grouped into two themes: human- data interaction (data and people), and facilitating technologies (policy, access, and standards). Below we describe each component of a PSDI. PSDI is both a theory defining what elements must be addressed to support effective spatial data use and the realization of said theory.

Users: A PSDI seeks to remove the burden of data processing from the user, and to improve data access, discovery, and usage to support increased focus on the resulting science. ***Policies***: An effective PSDI requires policies to ensure that community standards support the collection and sharing of data, and to ensure longevity and evolution of infrastructural services. ***Standards***: The development, codification, and adoption of data formatting and delivery standards to support data interoperability and use in widely available tools are essential for ensuring data usability. ***Data Access***: Effective data access is embodied by the ability to efficiently discover, ascertain, and utilize spatial data. Fundamentaly, access mechanisms are dependent upon the use of standard, interoperable formats by data providers. Rapid technological advancement requires access mechanisms that are adaptable as standards and protocols change. ***Data***: Data can be divided into two broad categories: *foundational* and *framework*.. The former includes geodetic control, topography, and orthorectified images [1], and are essential as baseline data products across a range of scientific and decision-making processes. Framework data are those products of critical importance to a smaller subset of the research community. Framework data may be used for a specific scientific objective, such as geologic or thematic mapping of a planetary surface.

The Planetary Data System: The NASA-sponsored PDS is “the formal archive for the planetary sciences” [10], created to preserve and make available data from NASA missions to the planetary science community. The PDS is comprised of six federated discipline nodes (Atmospheres, Cartography and Imaging Science, Geosciences, Planetary Plasma Interactions, Ring-Moon Systems, Small Bodies) and two technical support nodes (Engineering, Navigation and Ancillary Information). PDS personnel work with planetary mission instrument teams and individual data providers to plan and implement ingestion of peer-reviewed archives that meet specific standards, using PDS-4 protocols and formats. These archives are then made available on a world-wide basis using web-based interfaces.

Usability as a Common Goal: The PDS primarily supports data preservation, integrity, and access, while PSDI emphasizes data integration and interoperability for improved discoverability and usability. Although both entities ultimately support adherence to standards to improve data usability, this goal is addressed in different, but complementary ways.

In addition to the fundamental role of PDS in capturing, archiving and serving planetary data, the PDS Roadmap Study Team [RST, 10] recognized the fundamental need for improved data usability as a requirement for supporting users of planetary data. The RST report focused initially (Section 3.3) on data usability from the perspective of data discoverability, including the integration of data with metadata to facilitate the development of deep understanding of the data characteristics. In this sense, the data are ***usable*** when the user can find and access data via searches of available, relevant metadata. The RST report [10] also identified usability as a form of long-term accessibility. In this case, data usability is addressed by the PDS practice of limiting the number and complexity of data formats and requiring tools to be flexible in supporting them (i.e., “the PDS archival file formats are simple to support across generations (human and technological) without requiring format migration to preserve usability” [10, p.31]). This PDS approach can be described as an engineering view of data usability, comparable to that which was widely adopted by first- and second-generation terrestrial spatial data infrastructures (SDIs) [11].

In contrast, the proposed PSDI addresses development of a third-generation SDI that is fundamentally more user-centric. Third-generation systems depend upon the existence of first- and second-generation SDIs (the PDS being one example), but shift the focus of usability from the technical (engineering) to the user. For example, the PSDI view of discoverability leverages extensive, tightly coupled metadata (a technical requirement) with spatially and semantically enabled search capabilities to ensure that users can find desired data, as well as understand the spatial accuracy, spatial efficacy, and value of a given data product or set. Data are then not necessarily organized by science discipline, but by semantic meaning and contextual linkages. The identification of *foundational* and *framework* data sets [1] as completely distinct from a specific science application is evidence of this separation. Enabling semantic usability necessitates the generation and use of higher-order data, controlled, coregistered, and interoperable products that may not be usable over the long term in an engineering context. The PSDI view is driven by the desire to focus on the requirement that data not require spatial expertise to be utilized. In other words, the spatial data should just work [2] and this will invariably require (re)processing data archived in the PDS to higher-order, spatial products with reported spatial efficacy.

The Relationship Between PSDI and PDS: The distinction in approaches to defining and addressing usability has become apparent within the planetary science community, as exemplified by the RST report discussions, and thus it requires clarification. The largely engineering versus user-centric approach to data usability effectively delineates the boundary between the PDS and PSDI. We note that high-quality solutions to address user needs for usability are being implemented currently within the PDS, including the map-based searches for data enabled by the PDS Cartography and Imaging Sciences Node's Planetary Image Atlas [e.g., 4] and Planetary Image Locator Tool (PILOT, [e.g., 4]), and the PDS Geosciences Node's Orbital Data Explorers [12]. These could be considered second generation PSDIs, with the goal of promoting cross-discipline and -mission data searches.

However, the goal of enhanced data usability should extend well beyond these capabilities to include explicit definitions of how data components and services should interact, what format standards should be utilized for high interoperability, what the lifetime of derived data products that support improved usability should be, and how infrastructural data services can be decoupled from interfaces. Thus the proposed PSDI seeks to extend the PDS data delivery services and reframe these issues from a user perspective to ensure that data become even more usable to support science and exploration needs. To support this need, planetary data should be made available in ways that remove the requirement for spatial and data processing expertise; this immediately means that long-term usable (engineering) formats and short term usable (end-user) formats sometimes will differ. Also, the PSDI approach de-emphasizes the need for archiving of software because usable data formats may evolve along with tool requirements, highlighting the need to maintain and archive the capability to move from long-term archived data to shorter term, user-focused data formats. While it may be desirable to capture and widely share the most usable data products, it may not be necessary for these to be archived by PDS for long-term preservation.

The PSDI Initiative: A PSDI framework has been developed to address user-centric and data interoperability issues directly in a manner similar to that used by the terrestrial community to transition from second to third generation SDIs. This framework supports disentanglement of the needs of a long-lived archive from the needs of a rapidly changing user landscape.

Conclusion: The PDS and the PSDI framework are complementary components of a mechanism to make raw data highly usable for end users while still maintaining long-term preservation. No single format, storage mechanism, or management structure can adequately support the myriad of competing goals inherent in both long- and short- to medium-term usability. Several of the current PDS data services serve as foundational first and second generation PSDIs from which user-centric, third generation PSDIs can be developed. The implementation of PSDIs will be dependent upon the PDS for lower order data products and long term availability. The PDS benefits from PSDIs providing the shorter term spatially enabled usability that the planetary science community is requesting through the creation of higher order, interoperable, spatially enabled data products and services that are flexibly available in a rapidly changing technical and standards compliance environment. Existing PDS efforts (described above as 2nd generation PSDIs) and future efforts by PDS or non-PDS entities can fulfill the vision of a user-centric planetary spatial data infrastructure composed of data processing services to delivery higher-order spatial data, standards compliant map services that are usable by multiple clients, and semantic search capabilities for improved data discovery.

References: [1] Laura et. al. (2017) *ISPRS Int. J. Geo-Information.* [2] Laura et al. (2016) PSV 2050. Abs.# 8110. [3] Gaddis and Arvidson, (2017) *AGU* Abs.# 225497. [4] Gaddis et al. (2017) *3rd PDW* Abs.# 7124. [5] Laura et. al. (2017) *VEXAG* Abs.# 8012. [6] Radebaugh et al. (2017) *LEAG* Abs.# 5053. [7] Bland et. al. (2017) *OPAG* (informal pres.). [8] Exec. Order 12906. (1994) OMB. [9] Rajabifard et al., (2001) SDI Concepts. [10] PDS Roadmap Study for 2017 – 2026. [11] McLaughlin and Nichols (1994), *J. Survey Engin.* 120, 2. [12] Wang et al., (2017) 3rd *PDW*, Abs.# 7026.

PLANETARY SURFACE VISUALIZATION AND ANALYTICS. E. Law[1], Solar System Treks Project Team[1], [1]Jet Propoulsion Laboratory, Califronia Institute of Technology.

Introduction: The Solar System Treks Project is a collaborative project led by Solar System Exploration Research Virtual Institute (SSERVI) at NASA's Ames Research Center. JPL leads the engineering and implementation with USGS as the primary contributor providing valued-added data product from various missions. The project has developed three operational web-based portals: Moon Trek [1], Mars Trek [2], and Vesta Trek [3] providing a suite of interactive visualization and analysis tools to enable users including mission planners, scientists and general public to access large amount of mapped lunar, Mars and Vesta data products based on data collected by past and current lunar missions, Mars missions and from the Dawn mission.

The portals allow users to explore and measure the surface, zoom in and out of surface of the planetary bodies. The interactive maps are provided with different overlay options that provide details including visualization of various types of data (e.g., topography, mineralogy, abundance of elements and geology etc). These maps are value-added products based on data available from the Planetary Data System (PDS) [4]. The portals also provide 3-D printer-exportable topography so users can print physical models of the Moon's, Mars' and Vesta's surface. In addition, standards keyboard gaming controls are available to maneuver a first-person flyover view across the surface of these planetary bodies.

We will give an overview and direction of the project, including highlights of the operational portals and portals in work, as well as demonstration of their key features.

References:

[1] https://moontrek.jpl.nasa.gov/
[2] https://marstrek.jpl.nasa.gov/
[3] https://vestatrek.jpl.nasa.gov/
[4] https://pds.nasa.gov/

MoonDB - A Data System for Analytical Data of Lunar Samples

K. Lehnert, P. Ji, M. Cai, C. Evans, R. Zeigler

MoonDB is a data system that synthesizes analytical data acquired on lunar samples and makes them easily accessible online as a single data product with human and machine-readable interfaces that integrate the data into digital research data infrastructure. MoonDB preserves and restores data that have been scattered across the scientific literature, in online PDF documents, and in private files, having made discovery and reuse of these data difficult at best. MoonDB also serves as a home for new data generated on lunar samples, integrating new measurements seamlessly with the historical data. MoonDB helps researchers comply with US federal mandates for open access to research results.

MoonDB started as a collaboration between the Geoinformatics Research Group (GRG) at the Lamont-Doherty Earth Observatory, the Astromaterials Acquisition and Curation Office (AACO) of NASA's Johnson Space Center (hereafter JSC Curation), and senior members of the lunar science community. Over the past two years, the MoonDB project focused on (1) developing the MoonDB data system consisting of a relational database to store data and metadata, and a web application that provides users with web-based software tools to browse, search, and retrieve data; (2) summarizing references for lunar geochemistry and petrology from all relevant published sources and the Apollo Sample Compendium into the MoonDB Reference Catalog; (3) compiling geochemical and petrological data and metadata from these references and from datasets contributed by researchers into data templates and ingest these into the MoonDB database; and (4) supporting lunar scientists with preparing their unpublished geochemical data for ingestion into MoonDB. To encourage contribution of new data, the MoonDB web site features a service for investigators to contribute their data. Submitted data are deposited and published in the EarthChem Library, the geochemistry data repository recommended by publishers (e.g. Nature Scientific Data), where data are assigned a persistent unique identifier (DOI) so that they can be formally cited to give credit to authors, and linked to publications.

Figure 1 provides an overview of MoonDB's architecture and software stack. The system consists of the storage database (PostgreSQL), reporting database (ElasticSearch), the MoonDB search application (Python & JavaScript), and the MoonDB API, which serves as the bridge between the ElasticSearch index and the search application and can also be used by external systems to access MoonDB data. Data are loaded into the SQL database via the MoonDBLoader web application.

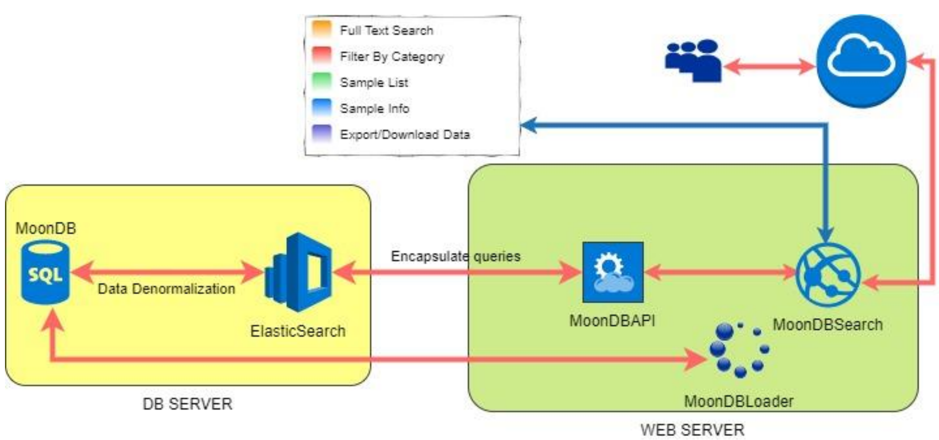


Figure 1 Architecture of the MoonDB Data System

The MoonDB database uses a customized version of the Observation Data Model (ODM2) relational database schema to store data and metadata of included datasets. ODM2 is a community information model designed to support a wide variety of feature-based Earth observations derived from sensors and samples, and to improve the capture, sharing, and preservation of these data (Horsburgh et al. 2016; Hsu et al. 2017). ODM2 implements concepts of the Observations and Measurements (O&M) standard (Cox 2007). O&M is one of the core standards in the OGC (Open Geospatial Consortium) Sensor Web Enablement suite. The use of ODM2 gives MoonDB a flexible and scalable structure to store data and metadata for a diverse range of samples and analytical data and makes it compatible with the new EarthChem synthesis database, which is also built on ODM2.

The MoonDB search application provides the tools for users to explore the content of the MoonDB database, select the samples and data they are interested in (e.g., specific chemical parameters, analyzed material such as whole rock or minerals), view and browse them, and download them in a useful format. Users first choose samples based on attributes such as sample type (lithology), geographic feature, data availability, and reference. They can then select chemical parameters to create a customized dataset containing a set of analytical data for a specific sample or group of samples, retrieved from separately published datasets. The custom dataset can be viewed online and downloaded in different formats. The MoonDB search interface features free-text and structured, faceted searches to find, filter, and explore data stored in MoonDB, and a download option for .csv file.

The MoonDB API (Application Programming Interface) is the bridge between the database and web applications. It defines a set of Hypertext Transfer Protocol (HTTP) request messages and the structure of response messages. The API feeds the HTML5 web component of the MoonDB search application. The API supports any third party uses or applications to retrieve data from MoonDB, exposing queries that return data for samples, features, datasets, and authors. A first set of HTML5 web components that interested data facilities can embed in their web pages have been developed. These web components allow access to the sample metadata and analytical results in MoonDB from data systems that scientists may use to look for lunar data, e.g. the Lunar Sample Database, MoonTrek, the Analyst's Notebook, and other planetary data systems.

During its next development phase, the MoonDB project will add geochronological data and lunar meteorite data to the synthesis and advance integration with the Planetary Data System by developing PDS4-compliant versions of the MoonDB data as a lunar sample bundle with table, context image and document collections with XML labels and archive these with the PDS Cartography and Imaging Sciences Node (IMG).

Cox, S.J.D., Ed., (2007), Observations and Measurements – Part 1 – Observation Schema: Open Geospatial Consortium. Wayland, MA., OGC doc. no. 07-002r1, 73 p.

Horsburgh, J. S., Aufdenkampe, A. K., Mayorga, E., Lehnert, K. A., Hsu, L., Song, L., Spackman Jones, A., Damiano, S. G., Tarboton, D. G., Valentine, D., Zaslavsky, I., Whitenack, T., (2016), Observations Data Model 2: A community information model for spatially discrete Earth observations. Environmental Modelling & Software, 79, 55-74, doi: 10.1016/j.envsoft.2016.01.010.

Hsu, L., Mayorga, E., Horsburgh, J. S., Carter, M. R., Lehnert, K. A., Brantley, S. L., (2017), Enhancing Interoperability and Capabilities of Earth Science Data using the Observations Data Model 2 (ODM2). Data Science Journal, 16(4), 1-16, doi:10.5334/dsj-2017-004.

PDS4 DATA WITHIN THE PSA – A CROSS-MISSION AND CROSS-DISCIPLINE APPROACH TO A PDS4 ARCHIVE. T. L. Lim[1] , S. Martinez[1], D. Coia[1], I. Barbarisi[1], M. Barthelemy[1], S. Besse[1], D. Fraga Agudo[1], E. Grotheer[1], D. Heather[1], and C. Vallat[1], [1]European Space Astronomy Centre, Camino Bajo del Castillo, s/n., Urb. Villafranca del Castillo, 28692 Villanueva de la Cañada, Madrid, tlim@sciops.esa.int.

Introduction: The European Space Agency Planetary Science Archive (PSA)[1] is a repository for all European Planetary Science Missions. It currently hosts data from all missions since Giotto and continues to be populated currently with Rosetta, Mars Express and ExoMars 2016 data. ExoMars 2016 is the first mission in operations to use the PDS4 standard and the use of PDS4 in the PSA has been primarily a cross-mission development shared between the ExoMars 2016 and Bepi-Columbo mission teams. Ongoing development now also incudes the Exomars Rover Surface Platform (RSP) and JUICE mission teams.

The introduction of PDS4 provided an opportunity to produce a new PSA[2] which features a new user interface and many improvements to the database. To maximise the cross-mission use of data, the efficiency of the PSA design and the user experience, it was agreed to standardise the way that PDS4 missions in the PSA design their data bundles and structures. Addtionally PSA schema and schematron have been developed for housing attributes the PSA believe are applicable as cross-mission attributes.

This paper describes the implementation of PDS4 in the PSA both in terms of the data structure rules adopted and the PSA schema along with information on why the choices were made. The implementation for the ExoMars 2016 is described as an example of a remote sensing mission. Since its launch several challenges have been faced in developing the ExoMars 2016 database and these challenges will also be described. During 2017 the development of the Exomars RSP archive data design has commenced and the initial design for the Rover data is also illustrated.

References:

[1] Besse, S., et al. PSS, (2017)
[2] Macfarlane, A. J., et al. PSS, (2017)

CHARACTERIZING THE SPECTRAL RADIANCE OF LUNAR PERMANENTLY SHADOWED REGIONS P. Mahanti[1], H.M.Brown[1], M.S.Robinson[1], A.K. Boyd[1], D.Humm[2], A. Awumah[1], [1]LROC Science Operation Center, School of Earth and Space Exploration, Arizona State University, Tempe, Arizona (pmahanti.lroc@gmail.com);[2]Space Instrument Calibration Consulting, Annapolis, Maryland.

Introduction: The low obliquity of the Moon (1.54°) results in permanently shadowed regions (PSRs) in topographic lows at high latitudes [1, 2] (Figure 1). PSRs typically exist in the lower reaches of host craters while the upper reaches receive periods of solar illumination. Some permanent shadows have extremely low temperatures (measured mid-day brightness temperatures of 29 °K) and thus volatile molecules deposited from various sources (e.g. comet or asteroid impacts) can remain cold trapped [3]. Mapping the location, abundance, and species of volatiles within PSRs is a key goal of previous, ongoing, and upcoming lunar missions.

PSR temperatures (and cold-trap behavior) is controlled by scattered solar illumination bouncing off crater walls and other surrounding topography, emitted thermal energy from surrounding topography, and energy from the interior of the Moon [4]. By using topography based simulation of scattered light, the incoming flux at some PSRs (>200 m diameter) (irradiance; Wm^{-2} units) was characterized earlier [5] but a detailed observation based analysis characterizing the magnitudes of scattered solar illumination is absent. In this work we study the broadband visible (400−760 nm) spectral radiance (R; $Wm^{-2}sr^{-1}\mu m^{-1}$ units) within PSRs from scattered solar illumination.

PSR Imaging: The Lunar Reconnaissance Orbiter Camera (LROC) Narrow Angle Camera (NAC) obtains useful images of PSRs with long exposure observations during times of maximum secondary illumination [6]. The increased exposure time 1.01 ms to 35.3 ms) leads to elongated pixels in the along-track direction, thus pixel scale is increased compared to typical NAC images (10–40 m/pixel vs. 0.5 m/pixel). Typically, the 12-bit (0-4095) NAC values have sufficient dynamic range to capture and reveal landforms in both illuminated and shadowed terrains. A raw NAC image is converted to a radiance calibrated NAC image by using the USGS ISIS LRONACCAL module[7, 8]

The acquisition of NAC PSR observations has been refined over several campaigns to optimize the trade-off between signal-to-noise ratio (SNR) and pixel scale, resulting in a comprehensive dataset [9]. A total of 6,199 individual NAC PSR observations have been collected since the start of the mission through 1 October 2017. The NAC PSR images reviewed in this study include a total of 4,368 unique NAC observations that overlap PSRs >10 km^2, ranging from 81°N,S to 90°N,S. PSR boundaries are identified from LOLA PSR shapefile [5] which is used to crop NAC images to PSRs.

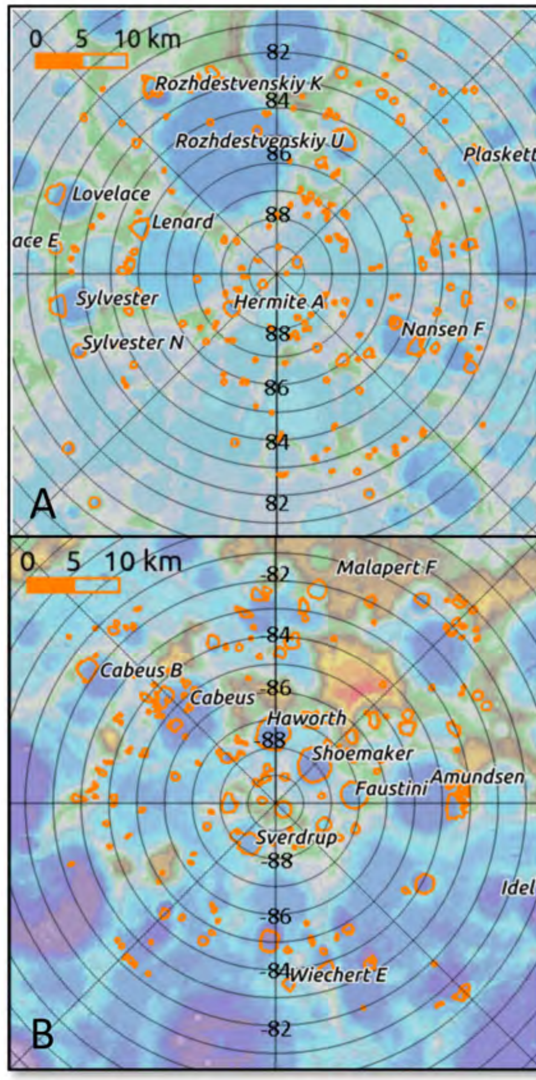


Figure 1: Polar stereographic views of the color shaded relief from the GLD 100, with LOLA PSR boundaries used in this study for the north pole (A) and south pole (B).

Computing spectral statistics: A total of 5,940 cropped PSR regions (multiple NAC images and multiple PSRs) were used to derive the radiance statistics within the PSR masks. For each PSR and overlapping image combination, the median (50% signal value from CDF), lower-limit (LL; 2% signal value) and upper-limit (UL; 98% signal value) statistics were obtained (Figure 2). These radiance statistics were used to compute the radiance cumulative distribution func-

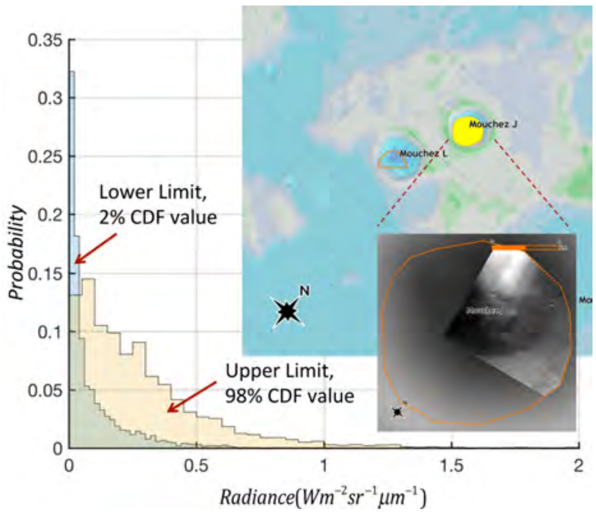


Figure 2: Distribution of Radiance UL (98%) and LL (2%) statistics. Inset shows an example of PSR outline (orange) partly imaged by a cropped NAC image.

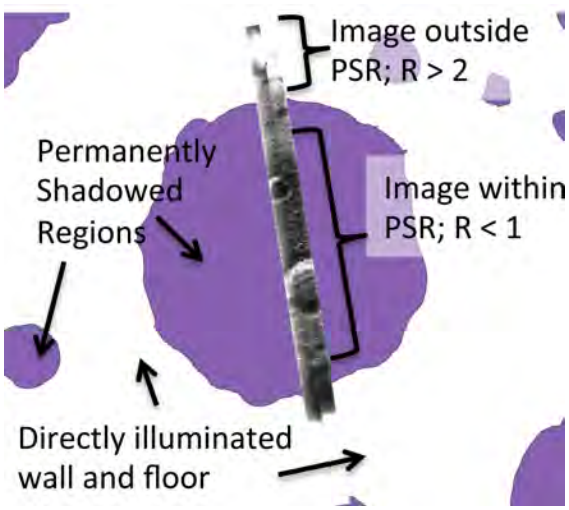


Figure 3: Radiance values on and outside the PSR within Shackleton crater.

tions (CDF). The value of the radiance signal depends on many factors. The most important ones are: (1) Solar position (latitude and longitude), (2) the angle of the orbital plane of LRO and the solar vector (β), (3) PSR latitude and longitude, and (4) Exposure time.

Results and Discussions: The median values of the 2% and 98% PSR scene radiance statistics are 0.04 and 0.2 $Wm^{-2}sr^{-1}\mu m^{-1}$ respectively, indicating the large number of pixels in most PSR images have $R < 0.2$. The maximum UL and LL values were 2.7 and 1.1 $Wm^{-2}sr^{-1}\mu m^{-1}$. The median radiance value from the NAC was 0.1 $Wm^{-2}sr^{-1}\mu m^{-1}$. Approximately 85% of the images have a dynamic range [2%, 98%] limits of 0.01 and 2 $Wm^{-2}sr^{-1}\mu m^{-1}$ and $Rvalues >$ 2 occur outside the PSRs (Figure 3). North pole PSRs

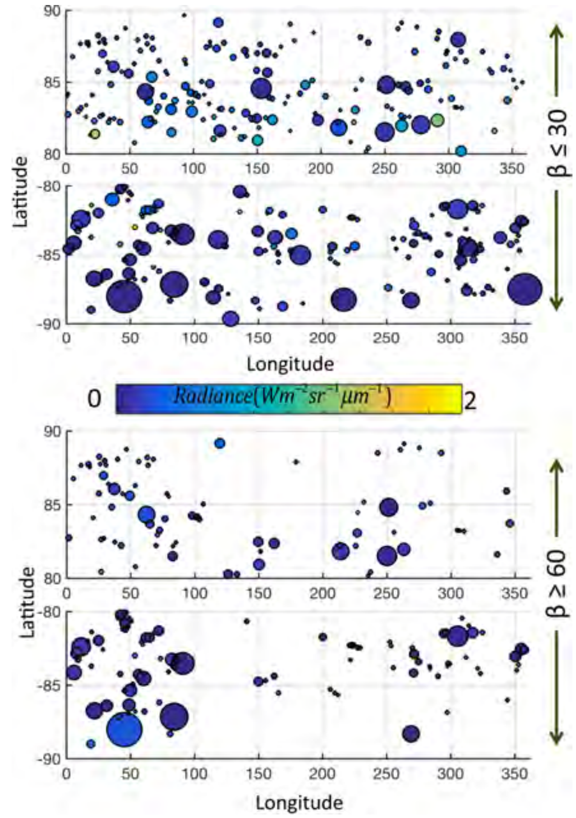


Figure 4: Radiance (UL statistics) at PSRs when images were acquired at high and low beta.

are typically smaller than south pole PSRs (median values of $25km^2$ and $77km^2$ for PSRs larger than $10km^2$) and are brighter as a group than south pole PSRs (Figure 4). When images are acquired at more acute β angles the disparity increases (UL values of 0.36 vs 0.15).

Conclusion: Imaging PSR scenes is challenging due to the minimal illumination at the PSRs (close to or below the detector limit) proximal to bright targets (illuminated terrain). The sensitivity and large dynamic range of the NAC helps PSR imaging, although at lower SNR and resolution compared to nominal NAC images. PSR scene radiance statistics computed in this work takes us a step further in being able to distinguish increased radiance levels from future PSR images of higher resolution and signal-to-noise ratio.

References: [1] H. UREY (1951) *Press Nueva Harven.* [2] K. Watson, et al. (1961) *Journal of Geophysical Research* 66(9):3033. [3] D. A. Paige, et al. (2010) *science* 330(6003):479. [4] A. R. Vasavada, et al. (1999) *Icarus* 141(2):179. [5] E. Mazarico, et al. (2011) *Icarus* 211(2):1066. [6] S. Koeber, et al. (2013) in *Lunar and Planetary Science Conference* vol. 44 2588. [7] L. Keszthelyi, et al. (2013) in *Lunar and Planetary Science Conference* vol. 44 2546. [8] J. Anderson, et al. (2004) in *Lunar and Planetary Science Conference* vol. 35. [9] S. Koeber, et al. (2014) in *Lunar and Planetary Science Conference* vol. 45 2811.

CLASSIFICATION OF SMALL LUNAR CRATER MORPHOLOGICAL STATE BY DEEP LEARNING P. Mahanti[1], K.Lanjewar[1], T.Thompson[1], LROC Team[1], [1]LROC Science Operation Center, School of Earth and Space Exploration, Arizona State University, Tempe, Arizona (pmahanti.lroc@gmail.com)

Introduction: The Moon is ideal for studying impact-cratering events that modify planetary surface morphology and lunar craters degrade over time such that their morphology changes as craters age. The study of impact crater formation and degradation over time (in the absence of strong weathering agents) provide valuable insights into the current meteoritic impact rates and surface target properties at scales relatable to future human exploration efforts. The rate of degradation can be estimated by the analysis of crater shapes at different morphological states (from fresh to degraded) over an exhaustive size-limited population of craters[1] within a specific geologic unit. However, morphological classification of a large population of craters, which is the first step in any such analysis involves a visual classification exercise which is repetitive, strenuous and subject to human error inclusions. Ultimately, this extremely time/resource-consuming step may significantly delay and even deter scientific analysis that could be achieved from morphologically classified groups. By automating the repetitive exercise of visual crater classification, consistent, efficient classification can be achieved while analysis time is freed-up for the user. Successful automation can also simplify the classification and scientific analysis of larger, more diverse populations with little loss in time (at the expense of computational power). We explore a novel deep learning[2] based classification strategy in this work to perform a binary classification task in planetary geology (whether a small lunar crater (diameter $< 300m$) is fresh or degraded) in this work. Specifically, a deep convolutional neural network (CNN) machine learning framework is utilized.

Deep Learning Method: TensorFlow[3] an open-source library for machine learning algorithms and Inception-v3, a deep learning model pre-trained with ImageNet[4] (an academic benchmark for computer vision) was used for this work. The pre-trained model Inception-v3 can classify the ImageNet dataset of 150,000 images into 1000 classes with minimal error, details can be found in [4]. ImageNet does not include lunar crater images, thus the two classes of lunar craters (fresh and degraded) are new to the pre-trained model. Since re-training the full Inception-v3 model from scratch is highly GPU intensive (may take weeks) we use transfer learning technique to retrain for new classes by using the existing weights (within the CNN) for known classes[5]. Our hypothesis is that by training only the final layer from scratch, reasonable classification performance for a large number of craters

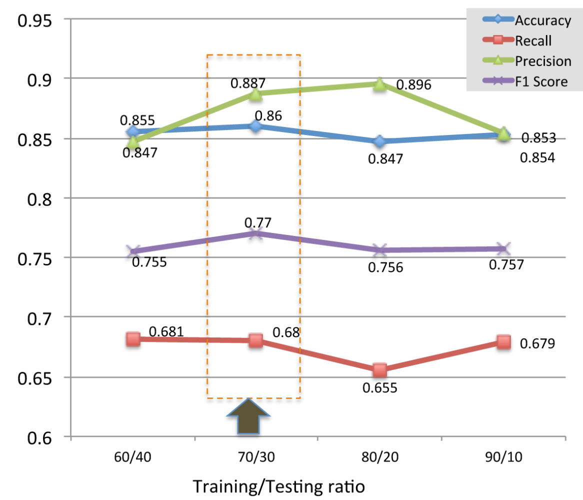


Figure 1: Selection of training to testing ratio of available observations

(> 1000), could be achieved for our work, in as little as thirty minutes on a laptop without GPU.

Training/testing images: Our testing and training dataset consisted of 5,569 pre-classified reflectance images of small craters (single, centered crater per image) acquired at 40 to 60 degrees incidence angle at the Apollo 16 and 17 landing sites. Image rasters were input as 8-bit JPEG files and had a resolution of 5 m/pixel.

Performance metrics: Performance of the trained model was quantitatively obtained from 5 metrics (1) Accuracy (ratio of correct classifications to the total observations), Recall (ratio of correctly classified fresh craters to actual number of fresh craters), Precision (ratio of correctly classified fresh craters to the total predicted fresh craters), F1-Score (weighted average of Precision and Recall) and MCC (Matthews correlation coefficient; a correlation coefficient between the observed and predicted binary classifications.

The training vs testing ratio (number of observations) was established by conducting a series of tests with different ratios (Figure 1). The 70/30 ratio (training to testing) was adopted for the final training of the model (optimal values of performance metrics).

For the final results the model was trained and tested 100 times to evaluate the consistency of model performance for each crater in the test set (n = 1670). Only unambiguous decision (no conflicting decision in 100 trials) was used to obtain the final percentages of fresh and degraded craters in the test set (Table 1). A crater was tagged un-decided if one or more conflicting deci-

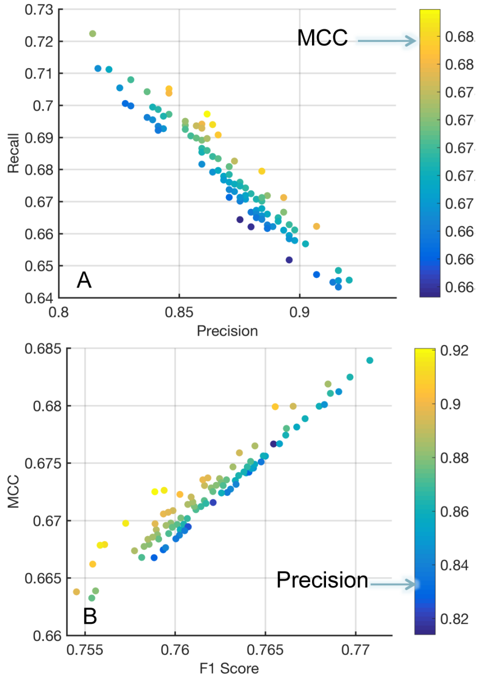


Figure 2: A: Inverse relation of Recall and Precision, B: Positive linear relationship of MCC and F1-Score, corresponding MCC scores are slightly lower

sion (fresh or degraded) was obtained during the 100 trials. Note that the confusion matrix represents 100 test runs also, and the average values of the computed entries (true fresh, false fresh, true degraded, false degraded) are used to compute the final performance metrics (Table 2).

Results and Discussions: The deep learning based method, achieved high values of overall accuracy (87%). In addition, the MCC value is close to 0.7 (high predictive correctness boundary [6]) showing that overall model performance is very good, even with unequal true-class sizes in the training and testing data. A large precision value ensures that more than 86% of craters identified to be fresh will be fresh, such that a trained deep learning model can successfully identify and classify fresh craters without supervision. Desired morphological properties can thus be further studied for only fresh craters from a large group of craters with minimal delay. We note however, the recall value is moderate to high, meaning that a portion of fresh craters will remain unidentified with our current model or training approach.

The classified craters in both Apollo 16 and 17 sites were further grouped to analyze performance for individual sites. The deep learning classification obtained fresh to degraded percentages similar to the manual classification (ground truth). The implication is that starting from a random sampling of nearly equal number of crater images at the two sites, the deep learning based classification can differentiate between the two sites as Apollo 17 being more degraded that Apollo 16, an observation obtained in an earlier work[1]. The distribution of depth-to-diameter ratio (depth obtained from digital terrain models) for fresh and degraded craters closely tracks (Figure 3) the true (manual classification) distribution curve implies an unbiased classification performance (if one class is preferentially identified correctly the distribution shapes will not match).

Table 1: Deep learning based classification results

Classification Method	**Manual**		**Deep Learning**	
Apollo Site	16	17	16	17
Total No. of Craters	849	821	849	821
%Fresh	31	21	34	24
%Degraded	69	79	56	67
%Undecided	0	0	10	9
Ratio (Degraded /Fresh)	2.2	3.7	1.7	2.8

Table 2: Performance Metrics

Accuracy	**Recall**	**Precision**	**F1 Score**	**MCC**
0.86	0.67	0.87	0.76	0.67

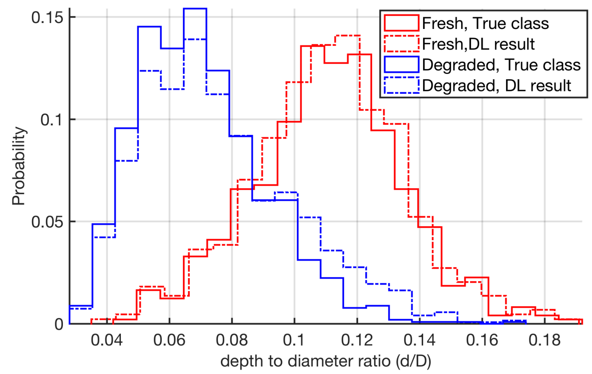


Figure 3: Distribution of depth-to-diameter ratio from visual and deep learning classification

Conclusion and Future Work: Deep learning based methods can drastically minimize the time for classifying lunar crater morphological states from crater images, boosting available time for class-specific scientific analysis, e.g. estimation of degradation rates. In our future work, a larger training dataset and a modified learning strategy will be used.

References: [1] P. Mahanti, et al. (2018) *Icarus* 299:475. [2] Y. LeCun, et al. (2015) *Nature* 521(7553):436. [3] M. Abadi, et al. (2016) *arXiv preprint arXiv:160304467*. [4] C. Szegedy, et al. (2016) in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2818–2826. [5] J. Donahue, et al. (2014) in *International conference on machine learning* 647–655. [6] D. M. Powers (2011) .

FITS AND PDS4: PLANETARY SURFACE DATA INTEROPERABILITY MADE EASIER. C. Marmo[1], T.M. Hare[2], S. Erard[3], B. Cecconi[3], M. Minin[4], A. P. Rossi[4], F. Costard[1] and F. Schmidt[1], [1]GEOPS, Univ. Paris-Sud, CNRS, Univ. Paris-Saclay, Rue du Belvédère, Bât. 509, 91405 Orsay, France, chiara.marmo@u-psud.fr, [2]U. S. Geological Survey, Astrogeology Science Center, Flagstaff, AZ, [3]LESIA, Observatoire de Paris, PSL Research University, CNRS, Sorbonne Universités, UPMC Univ. Paris 06, Univ. Paris Diderot, Sorbonne Paris Cité, Meudon, France, [4]Jacobs University, Bremen, Germany.

Introduction: Planetary science is a vast field of investigation that brings together several research communities (geologists, astronomers, physicists, geochemists, etc.), and produces an impressively growing amount of heterogeneous data. Interoperability and openness of data formats and processing techniques are becoming a necessity, to avoid the risk of being unable to efficiently extract scientific information from the data, and to guarantee the reproducibility of the scientific results.

This abstract will describe how Flexible Image Transport System (FITS) [1, 2] can be used in planetary surface investigations, and how its metadata can easily be inserted in the PDS 4 [3, 4] metadata distribution model.

FITS for Planetary Surfaces: FITS is an open digital standard, defined by the astronomical scientific community for data acquisition and archiving in astronomical observatories back in the late 70's, and is used for spatial telescope data too. FITS is one of the standard formats in the Virtual Observatory (VO) and it is compatible with PDS (version 3 and 4) archiving specifications. It is supported by a large number of open libraries and software tools, including the growing Astropy [5, 6] initiative, that provides an efficient and well documented framework for astronomical data reduction pipeline in Python.

The option to use FITS within the planetary domain is an opportunity to allow sharing of data across different domains and homogenize methods from acquisition, to visualization, while optimizing data processing.

FITS is already able to propose standard formatting for some data products quite common by now in planetary surface investigations. In particular, Multi-Extension FITS (MEF) schema proposes an easy way to store inhomogeneous digital information (reflectance, calibration data, vector table data, etc.) in the same file each with relative metadata, as well as multi-detector imagery (e.g. from HiRISE [7]) or hyperspectral cubes (e.g. from CRISM [8] or OMEGA [9]). FITS has been already chosen to distribute, e.g., Hayabusa [10] and some of the Dawn [11] data.

To be efficiently used in planetary surface investigations, FITS metadata must be extended in order to take into account the size of the reference body. In the framework of the VESPA [12, 13] component of the Europlanet 2020 project an extension to FITS metadata (GeoFITS) has been proposed [14].

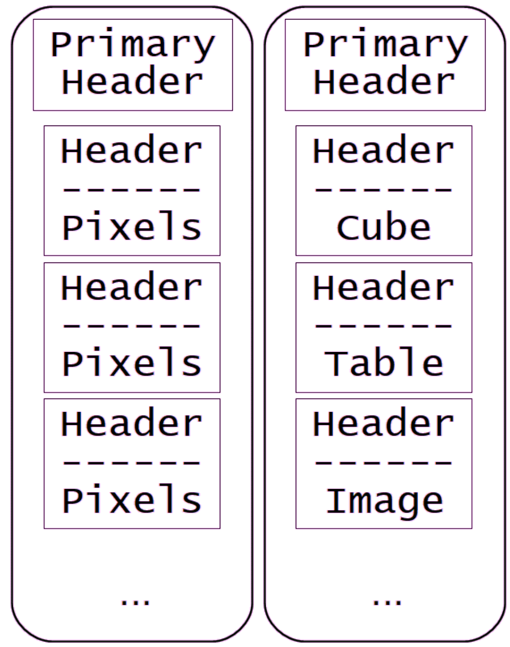


Fig. 1: Scheme of Multiple Digital Objects as MEF

Developments are in progress [15] in order to provide interoperability between the FITS format and geospatial applications commonly used by planetary surface research community by using the open source Geospatial Data Abstraction Library (GDAL) [16].

PDS4 evolution: PDS standard version 4 completely assumes its role of data modeling for archiving and distribution, detaching archiving metadata from data. This means that in PDS 4 archiving structure, labels and data products are always detached. PDS4 labels are written in eXtensible Markup Language (XML). XML was chosen because it is a widely used international standard for which a large amount of software already exists. In particular XML is used as exchange format in a wide number of on-line applications: PDS4 will simplify retrieving archive metadata by remote clients.

Like PDS3, PDS4 has a set of tags to support geospatial applications. Developments to provide interoperability to geospatial applications have already started [17, 18].

FITS to PDS4 dictionary: As previously outlined, PDS4 is meant to be an archiving standard, while FITS is proposed here to simplify and automatize processing logistics. PDS4 Classes and Attributes do not correspond literally to FITS keywords. The relation between them is highlighted in Table 1 as a way of example, to show how standard FITS keyword can be used to export a data product model in PDS4.

Table 1: Standard FITS keywords in the PDS4 archiving scheme

FITS	PDS4	Notes
BITPIX	Data_type	8 → UnsignedByte or UnsignedBitString 16 → SignedMSB2 32 → SignedMSB4 -32 → IEEE754MSBSingle -64 → IEEE754MSBDouble
BSCALE	Scaling_Factor	*Real*
BZERO	Value_Offset	*Real*
BUNIT	unit	*Short_String*
DATAMAX	valid_maximum	
DATAMIN	valid_minimum	
NAXIS	axes	axis_index_order = Last Index Fastest
DATE	creation_date_time	
DATE-OBS	start_date_time	
INSTRUME	Instrument	
TELESCOP	Instrument_Host	
OBJECT	Target	
ORIGIN	Institution_Name	
AUTHOR	author_list	
REFERENC	citation_information	

New FITS keywords proposed in the GeoFITS extension will be used to fill the PDS4 Cartography discipline namespace. See Table 2 and 3 for simple translations to the Coordinate_Representation and the Geodetic_Model classes.

Table 2: Filling Coordinate_Representation class with FITS metadata

PDS4	FITS
pixel_resolution_x	CD1_1 (in meters)
pixel_resolution_y	CD2_2 (in meters)
pixel_scale_x	1/CD1_1 (in degrees)
pixel_scale_y	1/CD2_2 (in degrees)

Table 3: Filling the Geodetic_Model class with FITS metadata

PDS4	FITS
latitude_type	WCSNAME
spheroid_name	WGCCRECS
semi_major_radius	A_RADIUS
semi_minor_radius	B_RADIUS
polar_radius	C_RADIUS
Longitude direction	CD1_1

References: [1] https://fits.gsfc.nasa.gov/. [2] Wells D. C. et al. (1981) Astronomy and Astrophysics Supplement, Vol. 44, P. 363. [3] https://pds.nasa.gov/pds4/about/. [4] https://pds.jpl.nasa.gov/pds4/doc/sr/current/StdRef_1.9.0.pdf. [5] http://www.astropy.org/. [6] Robitaille T. M. (2013) Astronomy & Astrophysics, Volume 558, id.A33. [7] McEwen A. S. et al. (2007) Journal of Geophysical Research, Volume 112, Issue E5. [8] Murchie S. et al. (2007) Journal of Geophysical Research, Volume 112, Issue E5. [9] Bibring, J.-P. et al. (2005), Science, doi:10.1126/science.1109509. [10] https://darts.isas.jaxa.jp/planet/project/hayabusa/data.html. [11] https://sbn.psi.edu/pds/archive/dawn.html. [12] http://europlanet-vespa.eu/. [13] Erard S. et al. (2017) Planet. Space Sci., doi:10.1016/j.pss.2017.05.013. [14] https://epn-vespa.github.io/geofits/. [15] Marmo et al. (2016) *LPS XLVII*, Abstract #1870. [16] http://www.gdal.org/gdal_vrttut.html. [17] Hare T. M. et al. (2017) PSS, doi:10.1016/j.pss.2017.04.004. [18] http://www.gdal.org/frmt_pds4.html

Acknowledgments: This work benefits from support of VESPA/Europlanet. The Europlanet 2020 Research Infrastructure is funded by the European Union under the Horizon 2020 research and innovation programme, grant agreement N° 654208.

NUCLEUSHUB.ORG: A PLATFORM FOR COLLABORATION AMONG ASTRONOMERS, NUCLEAR ASTROPHYSICISTS, AND PLANETARY SCIENTISTS. Bradley S. Meyer. Department of Physics and Astronomy, Clemson University, Clemson, SC 29634-0978, USA (mbradle@clemson.edu)

Introduction: The discovery of exoplanets has been one of the most exciting developments in all of science in the last few decades (e.g., [1]). The finding that planets around other stars are common, not rare, dramatically challenges our perspective on our Solar System and our place within the Universe. Scientifically, it directly links study of our own planetary system with study of planetary systems around other stars.

In a similar vein, presolar grains (e.g., [2]) and isotopic anomalies (e.g., [3]) in primitive meteorites are instances of the incomplete dissolution of the original building blocks of the Solar System. Those building blocks are interstellar dust grains, many of which are direct condensates from stellar outflows. As those dust grains traveled through the Galaxy, they sputtered and re-accreted, gradually homogenizing their isotopic abundances. Isotopic homogenization in the Galactic interstellar medium and in the early Solar System was not complete, however, and isotopic signatures in presolar grains and other primitive phases in meteorites record memory of the stellar processes that created them. A proper understanding of these effects thus requires connecting planetary science with the study of stellar evolution, element formation in stars, and Galactic chemical and isotopic evolution.

From these examples, it is evident that advances in planetary science will require increasing collaboration among astronomers, nuclear astrophysicists, and planetary scientists and cosmochemists. Such collaboration will benefit greatly from effective means of sharing information among these fields and, perhaps more importantly, of allowing the different fields to instruct each other in terminology, key ideas, and new discoveries.

A Hub for Collaboration: In order to facilitate collaboration among astronomers, nuclear astrophysicists, and planetary scientists, the author and collaborators are developing an experimental collaborative platform nucleushub.org. This site is built on HUBzero.org technology, which includes a powerful content management system to support scientific activities. Users are able to blog, participate in discussion groups, work together in projects, publish datasets and computational tools with Digital Object Identifiers (DOIs), and make these publications available for others to use.

A particularly powerful aspect of HUBzero is the ability to publish simulation and modeling tools that can be run on cloud resources. Middleware allows already existing tools to be deployed on the web with minimal modification. The Rappture toolkit is the infrastructure that allows developers to focus on their core algorithms when developing online simulation tools. The simplicity of development with Rappture means that students, even at the undergraduate level, can contribute useful tools to an operating HUB with a small amount of training.

Current Tools: The author's group is primarily interested in nuclear physics, nuclear astrophysics, stellar evolution, and Galactic chemical evolution and their implications for cosmochemistry. The group is developing a number of tools related to these topics and deploying them at **nucleushub.org** to illustrate key concepts. These tools include a nuclear partition function tool, a Solar abundances tool, and a simple Galactic chemical evolution simulator. The group hopes these tools will provide examples that other users can follow to contribute their own tools, and the author is actively recruiting collaborators around the world in a variety of fields to develop and contribute such tools. The hope is that **nucleushub.org** can evolve into a valuable platform linking planetary science to related fields of astronomy and nuclear astrophysics.

References: [1] Author A. B. and Author C. D. (1997) *JGR, 90,* 1151–1154. [1] Mayor M. and Queloz D. (1995) *Nature, 378,* 355-359. [2] Clayton D. D. and Nittler L. R. (2004) *Ann. Rev. Astron. Astrophys. 42,* 39-78. [3] Dauphas N. et al. (2004) *E&PSL, 226,* 465-475.

SCIENTIFIC HYBRID REALTIY ENVIRONMENTS (SHyRE): BRINGING FIELD WORK INTO THE LABORATORY.

M.J. Miller[1], T. Graff[1], K. Young[2], D. Coan[3], P. Whelley[4], J. Richardson[4], C. Knudson[4], J. Bleacher[4], W.B. Garry[4], F. Delgado[5], M. Noyes[5], P. Valle[5], J. Buffington[5], A. Abercromby[5], [1]Jacobs, NASA JSC, Houston, TX 77058 (matthew.j.miller-1@nasa.gov), [2]Jacobs/UTEP/NASA JSC, [3]Aerospace Corporation/NASA JSC, [4]NASA Goddard, [5]NASA JSC.

Introduction: The use of analog environments in preparing for future planetary surface exploration is key in ensuring we both understand the processes shaping other planetary surfaces as well as develop the technology, systems, and concepts of operations necessary to operate in these geologic environments [1], [2]. While conducting fieldwork and testing technology in relevant terrestrial field environments is crucial in this development, it is often the case that operational testing requires a time-intensive iterative process that is hampered by the rigorous conditions (e.g. terrain, weather, location, etc.) found in most field environments. Additionally, field deployments can be costly and must be scheduled months in advance, therefore limiting the testing opportunities required to investigate and compare science operational concepts to only once or twice per year.

To overcome these inherent challenges, SHyRE (Scientific Hybrid Reality Envirionments) is a Planetary Science and Technology from Analog Research (PSTAR) funded, multi-year campaign aimed at developing a scientifically-robust analog environment using a new and innovative hybrid reality (HR) setting that addresses these limitations and enables frequent operational testing and rapid protocol development. HR is unique in that operators not only work within a virtual environment, but physical objects, advanced tracking systems, and various other technologies (e.g. procedure assistant, voice recognition, torso/limb/finger tracking, etc.) are also incorporated to create a highly realistic and immersive simulated environment (Fig. 2). The application of this analog environment has immediate implications and opportunities to inform future planetary missions and science investigations by rapidly prototyping and testing new scientific instruments with relevant data processing activities (e.g. archiving and analysis) embedded within realistic/envisioned flight operational constraints.

The SHyRE program objectives are divided into two main categories: Technology development of the HR environment and Scientific Operations development of operational constraints consistent with crewed planetary surface exploration (e.g. planetary extravehicular activity). To facilitate these objectives, an Earth-based analog environment is leveraged to construct a realistic HR environment and preliminary concepts of operations are constructed from a combination of existing spaceflight architectures and prior analog studies.

This abstract briefly summarizes the HR environment development process as well as the preliminary EVA science operations envisioned within this digital setting with an emphasis on demonstrating development throughout this multi-year campaign.

Hybrid Reality Development: To produce a scientifically-relevant environment in HR, we selected the December 1974 (D1974) flow located in the SW rift zone at Kilauea Volcano, HI. This site was selected because 1) it serves as a realistic analog to Mars surface conditions and 2) has extensive existing prior field study data. Existing data sets were leveraged from multiple prior field deployments during the SSERVI RIS^{4}E (Solar System Exploration Research Virtual Institute; Remote, In Situ and Synchrotron Studes for Science and Exploration) program. These data sets consisted of field portable handheld x-ray fluorescence (hXRF), x-ray diffraction (XRD), light detection and ranging (LiDAR; for surface texture), and in-situ spectroscopy (VIS/NIR) [3]. To complete this suite of data sufficient for HR development, the SHyRE team performed one final field deployment to the D1974 flow where additional portable geochemical field instrument data using Laser Induced Breakdown Spectroscopy (LIBS) and additional XRF were collected in addition to supplemental LiDAR scans.

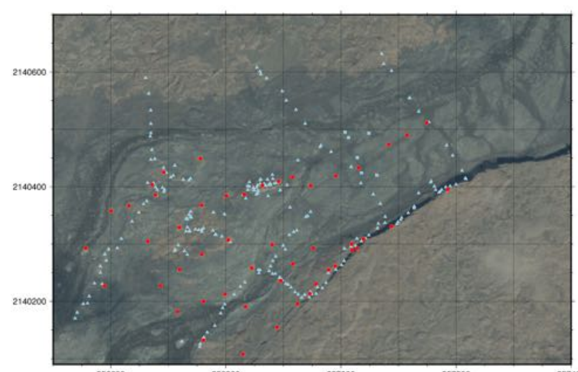


Figure 1. Geochemistry data (blue triangle) and LiDAR (red circle) GPS locations on the D1974 flow.

The SHyRE goal is to translate in situ instrument data taken from the D1974 flow and render the millimeter scale resolution of LiDAR data with geochemical signatures from the XRF, LIBS, and XRD instruments within the SHyRE HR environment, a process currently underway. As shown in Figure 2, the HR environment can already render Internation Space Station modules with handheld instruments such as the pistol grip tool (PGT).

Figure 2. NASA's HR lab showing a user viewing a photorealistic pistol grip tool (PGT) inside an HTC Vive headset, while manipulating a 3D printed PGT that allows for physical feedback.

Science Operations Development: Numerous prior NASA analog programs have already started to investigate crewed scientific planetary surface exploration architectures [4], [5]. SHyRE builds upon these prior tests and is being built instead to focus on how a crew will operate in the absence of real-time science backroom support (Fig. 3), as would likely be the case in the long communications latency (40+ mins round-trip) in Mars operations.

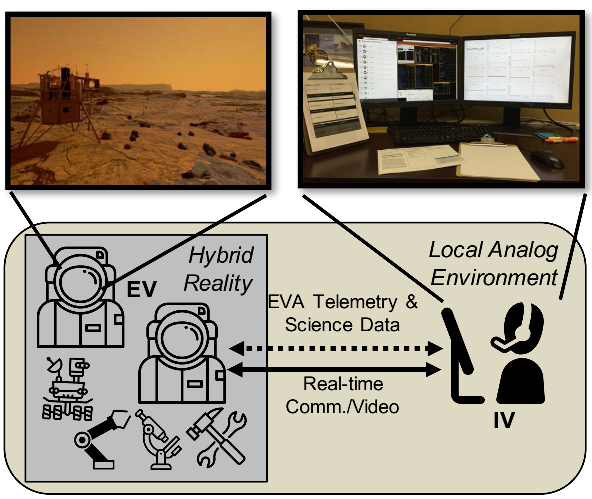


Figure 3. SHyRE concept of operations utilizing extravehicular (EV) crew in collaboration with an intravehicular (IV) crewmember to conduct EVA.

Under these operating conditions, the crew will be faced with a high volume of scientific instrumentation data coupled with the data pertaining to the vehicle and spacesuits. Development efforts are currently underway to prototype and render these instrument data products to both EV and IV crew for consumption during EVA execution to support scientific real-time decision-making. The HR environment affords an unprecedented level of simulation data acquisition regarding operationally relevant metrics (e.g. what the crews see, tool utilization, information transfer and interaction, etc) to objectively measure the execution of scientific operations. Prior work in both portable field instrument utilization [3] and EV and IV support system development [6] are being incorporated to support of this effort. In the coming year, SHyRE will investigate the relationship between crew workload and increased information flow (like that which would come with the addition of high-resolution field portable instruments) for scientifically-driven EVA.

Future Plans: By leveraging the HR setting and developing an environment with high scientific fidelity supported by multiple years of field data, the SHyRE Science Operations research will elevate the resolution and robustness of insights obtained from planetary analog operations research to an unprecedented level of detail. By conducting analog operations in the SHyRE HR environment, we can systematically control and vary relevant features (availability of scientific instrument data, etc.) of the concepts of operations in a repeatable way, something which has never been done before due to the inability to precisely duplicate test conditions in a traditional field environment. This notion of constructing testable, strictly-controlled environments aligns with and leverages experience from a multitude of established research efforts that examine the impact of new technologies and procedures in similarly complex environments (e.g. air traffic operations, maritime traffic management, and military command). Futhermore, SHyRE will be able to serve as a platform to develop, test and integrate the latest standards and process in scientific data management for future planetary exploration missions.

References: [1] W. B. Garry and J. E. Bleacher, (2011) Geological Society of America, vol. 483. [2] K. Young, et al., (2011) *Acta Astronautica*. vol. 90, no. 2. [3] K. E. Young, et al., (2016) *Applied Geochemistry*, vol. 72, no. C, [4] M. J. Miller, et al., (2017) AIAA SciTech #2017-1444, 2017 [5] K. H. Beaton, et al., (2017) IEEE Aerospace Conference [6] M. J. Miller, et al., (2017) *JCEDM*, vol. 11, no. 2.

Put Everything in a Database: Extremely Low Density Data Storage and Massive Relationality. C. C. Million, Million Concepts (PO Box 119, 141 Mary St, Lemont PA 16851, chase.million@gmail.com)

Introduction: The field of remote sensing is still largely locked into the use of rasterized data files. Very few GIS tools have any capacity to handle non-vector data that are not on a regular grid. This is possibly a holdover from a time when CCD (etc.) images were primarily regarded as digitized photographs, data storage volumes were limited and nonlinear access was slow, and processors were relatively less capable such that the linear algebraic optimizations enabled by regularly gridded data were necessary to perform many standard operations within a reasonable runtime.

Also, most existing pipelines for remote sensing data irrevocably sever the logical connection between data at different levels of processing. This is often, but not always, because reduced data at intermediate stages are unnecessarily interpolated or rasterized—particularly after geometric transformations—before serving as inputs to further processing steps. Most textbooks on remote sensing suggest that different types of interpolation be used—e.g. nearest-neighbor vs. bi-linear—depending on the type of analysis required in order to minimize the degradation in quality of information of interest. This suggests that there is only one interpolation step in preparing remote sensing data for analysis, but pipelines may have hundreds of interpolation steps, each one of which, by definition, artificially degrades the data.

The extent to which the data quality is degraded by repeated interpolation is very difficult to determine. The limitation in error propagation imposed by raster data formats may be a fundamental stumbling block in the problem of remote sensing data fusion to the extent that the data fusion problem is primarily caused by limited knowledge about uncertainties.

Recommendations: With the observation that data storage and processing are no longer practical limitations for most applications and informed by similar efforts to represent astronomical data flexibly in large databases (e.g. [1,2,3]), I have identified general principles for doing so that should maximize data quality, flexibility, and utility.

In particular:

1. Represent the raw observational data within the database such that each database row corresponds to the smallest and lowest-level meaningful datum. In a CCD image, for example, this datum would be the uncalibrated Data Number (DN) of a single detector pixel.
2. Link raw observational data (relationally) to metadata entries and tables.
3. Define database tables that record inputs and outputs to every step of processing.
    a. Optimally design processing steps to minimize information loss, particularly delaying interpolation or rasterization to as late as possible in processing.

Consequences: This representation of data offers several potential advantages:

- All operations on the data are, potentially, 100% lossless and reversible.
- Spectral and spatial errors can be precisely tracked on a per-pixel basis throughout all levels of processing.
- Information across all levels of processing is seamlessly available to incorporate into analysis. Questions such as, "What are the uncalibrated values of all pixels in the upper right quadrant of the detector that contain in an object that looks like a tree?" become trivial.
- Subsets of data can be precisely retrieved for further analysis without regard to the underlying data structure. This has been a limitation of high density data formats, in particular.
- Massively parallel data processing is possible for algorithms that require only subsets of the input data, for example pixel-scale vs. image-scale manipulations. A 10 megapixel image becomes 10 million single pixel images.
- Data processing is 100% reproducible—being entirely defined by the structure of the database—and easily modified at any processing step. Storage of the data at all intermediate processing steps within a database eliminates the need to reprocess up to that point.

I will describe the state of relevant technology and similar efforts as they relate to the feasibility of an extensible solution. I will also present the results of initial test implementations.

References: [1] Million, et al. ApJ, 833.2 (2016): 292. [2] Oosthoek, et al. (2012) Planetary Data Workshop abstract. [3] Stonebraker, et al. (2011) Springer Berlin Heidelberg.

Announcing a Community Effort to Create an Information Model for Research Software Archives. C. Million[1], A. Brazier[2] , T. King[3] , A. Hayes[2]. [1]Million Concepts (PO Box 119, 141 Mary St, Lemont, PA 16851; chase.million@gmail.com), [2]Cornell University, [3]UCLA

Summary: The planetary research community increasingly recognizes the value of archiving research software source code, executables, and operating environments for enabling other researchers to build upon prior work at the same time that some NASA programs require pulic release of such software for purposes of reproducibility. [1] But there is no PDS equivalent archive for planetary science research software at this time. Also, because software archiving is a relatively new practice, it has, as yet, no widely accepted standards. An effort has started within OpenPlanetary [2,3] to create recommendations and standards for the archiving of planetary science research software in the absence of an appropriate formal archive. The primary goal of this effort is to define an information model that can completely and effectively describe software data products and associated metadata and that is consistent with OAIS standards. [4] This will dramatically ease the ingestion of software into a formal archive when one becomes available. A prerequisite to this objective will be to establish what elements of data or metadata must be included as components of a software archive in order for that archive to effectively enable reproducible research. We invite interested members of the research community to join this activity by contacting any of this abstract's authors.

References: [1] Million, et al. (2017) Planetary Data Workshop abstract. [2] openplanetary.co [3] Manaud, et al. (2016) AAS: v.48. [4] ISO 14721:20012

ARCHIVE, ACCESS, AND SUPPLY OF SCIENTIFICALLY DERIVED DATA: A DATA MODEL FOR MULTI-PARAMETERIZED QUERYING WHERE SPECTRAL DATA BASE MEETS GIS-BASED MAPPING ARCHIVE. A. Naß[1], M. D'Amore[1], and J. Helbert [1]DLR, Institute for Planetary Research (Andrea.Nass@dlr.de, Mario.DAmore@dlr.de).

Introduction: Since the late 1950s a huge number of unmanned planetary missions were undertaken to explore our solar system. Since the 1990s, Europe has become highly active in planetary exploration with spacecraft contributions (e.g., Mars Express, Venus Express, Huygens probe, ExoMars, Rosetta), and the employment of dedicated mapping instruments. The data resulting from this robotic exploration and remote sensing varies in data type, resolution, and target. After different steps of pre-processing and correction, the released data are available for the community on different portals and archiving systems, e.g. *Planetary Data System* (PDS) or *Planetary Science Archive* (PSA).

One major usage for these data is *mapping*, i.e. the extraction and filtering of information by combining and visualizing different kind of base data. *Mapping* itself is conducted either for mission planning (e.g., identification of landing sites) or fundamental research (e.g., reconstruction of surface by multidimensional comparison of different base data (image data, spectral-/hyperspectral sensor data, radar images, and/or derived products like digital terrain model), identification of timing). The mapping results for mission planning are directly linked to and managed within particular mission teams. The valuable data and information derived from fundamental research - also describable as maps, diagrams, or analysis results - are mainly project-based and exclusively available in scientific papers. However, finding and accessing these valuable data to be used for further investigation is often not easy or downright impossible.

Therefore, one important question is how the derived mapping data described above can be archived comparably to the mission data, i.e. reusable, well-documented, and sustainable. A data archive is necessary, to enable further cross-links between different user groups and allows the reusability of already existing information and knowledge.

Thus, we discuss within this contribution:

Q1 How derived planetary scientific data like vector-based mapping, diagrams, and results of analysis can be archived, thus they could finally be used as additional base data for further investigations?

Q2 How different mission data (base data and derived products listed above) could be merged, to generate combined querying for the most efficient data and information handling?

Current Framework: Along with recent and upcoming missions also to Mercury (BepiColombo), the Outer Solar System moons (JUICE), and asteroids (NASA`s Dawn mission), systematic mapping of surfaces has received new impulses.

Since the late 1990s the scientific mapping community has started to use Geographic Information Systems (GIS) for planetary mapping. GIS frameworks are usually based on databases, which represent an ideal tool for generating, but also for archiving and storing spatial data - vector- as well as raster-based data.

To handle the two questions mentioned above, we build upon two developments, which are already established within the Institute of Planetary Science, DLR.

Part I: The Planetary Spectroscopy Laboratory (PSL) group at DLR joins the Participating Scientists for MESSENGER program for the Mercury Atmospheric and Surface Composition Spectrometer (MASCS) instrument, allowing access to the team data before the official release to PDS. MASCS have mapped Mercury surface in the 400–1145 nm wavelength range during orbital observations by the MESSENGER spacecraft. To overcome the dataset bulk size and fully exploit the information present in it, we developed a PostgreSQL/PostGIS distributed database. The DB contains the whole MASCS spectral dataset, around 4 Million single measurements as vector data, and user defined polygons. To explore possible relations between composition and spectral behavior, we have imported other dataset, like the elemental abundance maps derived from MESSENGER's X-Ray Spectrometer (XRS).

Part II: In the last years the Department of Planetary Geology, DLR established a GIS-based mapping archive (concept, and evaluation version) storing all different kind of derived vector-based mapping projects, which are conducting within different investigations. To enable and ensure a sustainable use of the derived data, two topics are treated: 1. Comparability and interoperability has been made possible by standard recommendations for visual, textual, and structural description of mapping data (e.g. [1], [2]). 2. Interoperability between users, information- and graphic systems is possible by templates for digital mapping and data bases (e.g. [3], [4]).

Therefore, this data base driven archive has to cover the requirement, (1) applicable for all known planetary bodies, (2) usable in the proprietary environment ArcGIS™ (ESRI), but also usable and accessible within independent and open GIS systems, like e.g. QGIS, (3) developed, or at least transferable, into a PostgreSQL/PostGIS driven data base structure, and last but not least, (4) the archived data should be available and replicable for future investigations.

One first implementation was conducted for the systematic mapping of Ceres (Dawn mission), is useable also outside the DLR, and was presented, e.g. [5].

Application: The current spatial intersection within *Part I* is a computation-heavy operation that is executed in the backend in period of low activity, typically at night. The current resulting features–measurements polygons intersection is stored in caching tables, allowing a quasi-live retrieve in GIS system from user perspective. The overhead in complexity is justified by the circumstance that the spatial query is executed only once, whereas the retrieving of the data could happen multiple times. Overall, despite the additional complexity and overhead to join different table, this approach optimizes the access time for spatial intersection. We are currently working on merging the GIS-based map archive to the PSL database to enable the data query for spectral data by the polygons, done within geologic/geomorphologic mapping projects.

Conclusion: The idea behind this contribution is to ingest the product of surface mapping done by experts (e.g., geomorphological or geological mapping), and intersect those features with the actual data, to extract spectral information in well know geological regions (Figure 1). The ingestion architecture expects a minimum set of feature to be defined by the user.

Three examples of this approach are: 1. the comparison spectral behavior with radial distance in more than 100 craters on the surface of Mercury (Figure 1, left) [6], 2. the identification of Olivine outcrops on the surface of Vesta via DAWN data analysis [7], 3. A general automated multi instrument mapping framework [8, 9].

The current approach shows that databases described as *Part I* and *Part II* are (1) theoretically transferable to any planetary body, e.g. from Moon, Mars, (2) through the spatial context all these data hold by nature, the two parts are combinable, this (3) enables an overarching and comparative research and analysis basis by multi-parameterized querying, and would (4) benefits the knowledge management and data/product usability for future missions and data.

Summary: An archive of already gained information supports the scientific community significantly by a constant rise of knowledge and understanding based on recent discussions within Information Science and Management, and Data Warehousing. An archiving structure and additional reference level of derived and already published data could easily be transferred to other scientific fields, and be linked to other planetary mission data, e.g. laser altimeter data [10].

References: [1] Naß, A. et al., AutoCarto, 2010, [2] Naß, A. et al., *PSS 59(11-12)*, p 1255-1264, 2011, [3] van Gasselt, S. & Naß, A., *PSS 59(11-12)*, p 1231-1242, 2011, [4] Nass, A., & van Gasselt, S., In: *Cartography from Pole to Pole*, Springer, p 261-270, 2013, [5] Naß, A., EPSC, #147-2, 2017, [6] P. D'Incecco et al., *PSS 132(32-56)*, 2016, [7] D'Amore, M. et al, 48th LPSC, 2017, [8] Domingue, D. et al., *Icarus*, 2017 (revised), [9] Domingue, D. et al., *Icarus*, 2017 (revised), [10] Stark, A. et al., this issue, 2018.

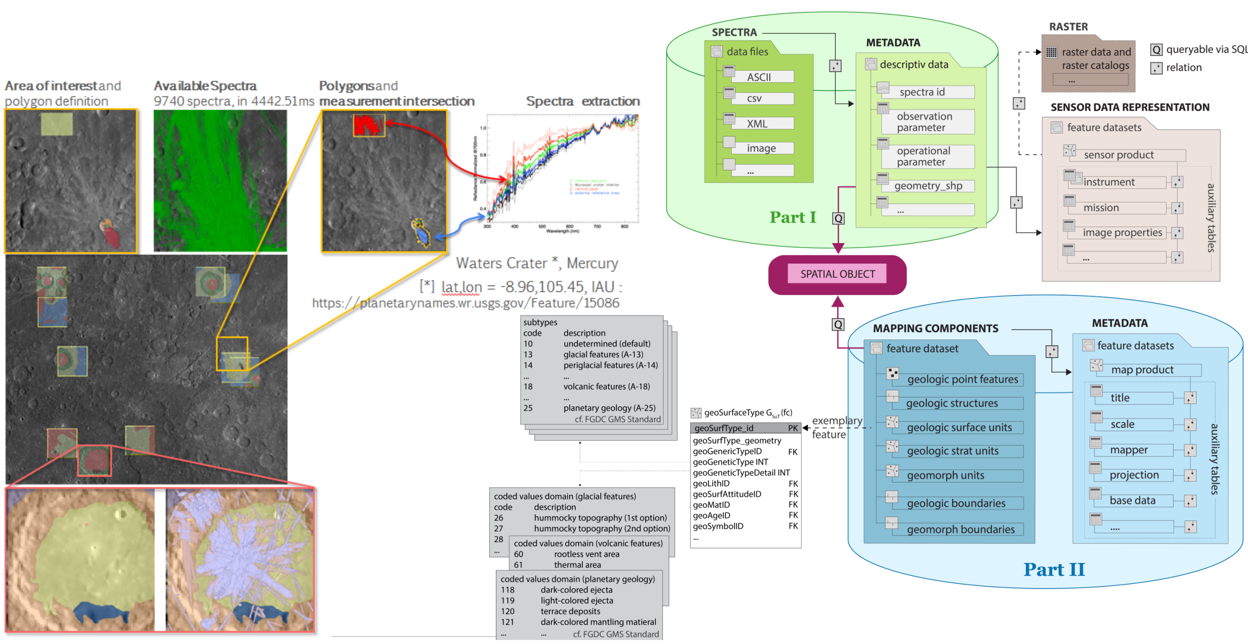


Figure 1 *left*: comparison spectral behavior with radial distance in more than 100 craters on the surface of Mercury [6], *right*: schematic and simplified model of *Part I* – spectral, and *Part II* – mapping database, inclusive metadata entries.

AUTOMATED DETECTION OF CRATERS IN MARTIAN SATELLITE IMAGERY USING CONVOLUTIONAL NEURAL NETWORKS. C. J. Norman[1], J. P. Paxman[1], G. K. Benedix[2], T. Tan[1], P. A. Bland[2], M. Towner[2] [1]Department of Mechanical Engineering, Curtin University, Kent Street, Bentley, Perth Western Australia, 6102, [2]Department of Applied Geology, Curtin University, Kent Street, Bentley, Perth Western Australia, 6102.

Introduction: Impact craters are structures formed when a meteoroid strikes the surface of a planetary body. The age of a surface can be estimated through analysis of crater frequencies, assuming random impact rates with known long-term averages [1]. Considerable research effort has been invested into developing automated techniques for the detection and counting of craters [2].

Approaches to Crater Detection Algorithms (CDAs) have included image analysis techniques such as edge detection and the Hough transform [3], though many in recent years have incorporated some form of machine learning, including neural network architectures.

We propose an automated machine learning solution to the crater detection and counting problem on Mars, involving a Convolutional Neural Network (CNN), with ground truth (training) data provided by the Robbins database of Martian craters with diameters over 1km [4].

Background: Convolutional neural networks have performed well at complex classification tasks, particularly in image recognition, where pixel proximity is exploited to solve the scaling problem which exists for conventional multilayer perceptron neural networks. The OverFeat CNN architecture is capable of object classification, localisation and detection. This architecture has been used to detect people in complex, crowded scenes [6], and was able to correctly detect overlapping and occluded examples, making it promising for use in a CDA – crater overlapping is similar to the occlusion problem.

Google Tensorflow is an open-source machine learning library based on data flow graphs. The advantage of a Tensor-flow based CDA is that the algorithm can be scaled to a distributed supercomputing environment to process large volumes of planetary data.

A supervised machine learning algorithm requires a ground truth database of examples labelled by experts. Robbins and Hynek [4] coordinated a community effort to catalogue Martian craters over the entire Martian surface with diameters above 1 km, currently numbering some 384,343 entries. We assumed the Robbins database to be consistent enough for the training of an algorithm, and sought to develop a CDA with the objective of generalising to craters smaller than 1km. Even among experts, a maximum variation of up to 45% has been noted in crater identifications [5]. A CDA may provide a more consistent, less subjective, and of course faster method of crater detection compared to expert analysis of images.

Approach: A CNN-based CDA was designed using TensorBox, an open-source object detection framework based on Google Tensorflow. Images from Mars Odyssey's Thermal Emission Imaging System (THEMIS) were used to generate a ground truth database using the Robbins identifications. The USGS Astrogeology Science Centre has released mosaic images combining individual THEMIS images into a larger map. Mosaics covering the entire equatorial latitude band of N°30 to S°30 were selected for analysis. Each mosaic image spans a region of approximately 2700 km by 1800 km. The mosaics were split into a total of 6387 tiles each with width 1280 pixels and height 960 pixels. The Robbins Database craters were mapped to each tile.

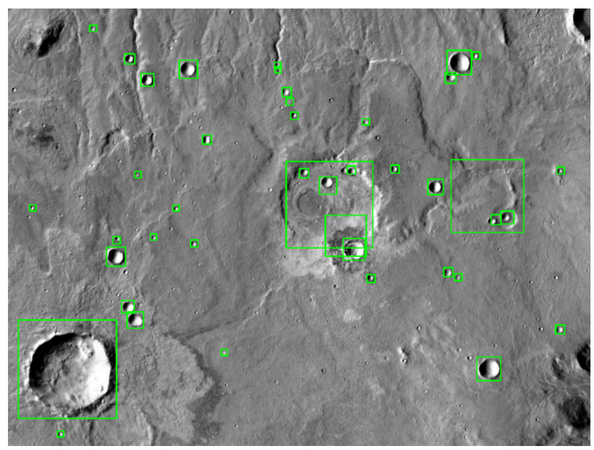


Figure 1: CDA crater detections

The CDA was trained for 20,000 iterations with a learning rate of 0.001. The neural network model that was used by the CDA was GoogLeNet-OverFeat, pretrained on the Imagenet dataset. Figure 1 shows the predictions made using the CDA. Preliminary results visually correlate with craters. The CDA is able to detect a broad range of craters that vary in size and appearance.

The algorithm was generalised to higher-resolution Context Camera (CTX) Martian satellite imagery to detect craters from approximately 100m in diameter in the vicinity of the Mojave crater. Figure 2 shows a crater size frequency distribution isochron generated through the automated workflow for comparison with earlier analysis by Werner et al. [7] based on manual counts. The crater frequencies are consistent with the results of Werner's analysis in the 0.1-1km range.

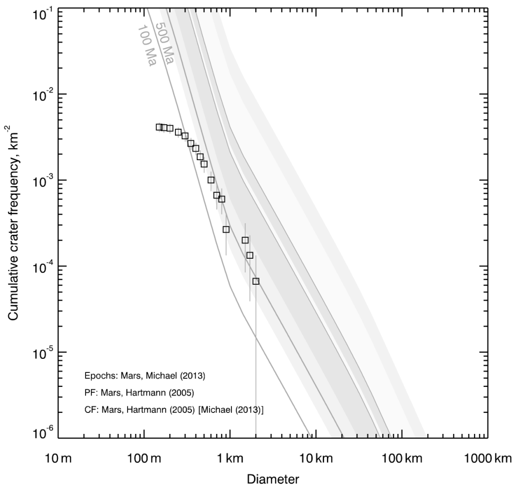


Figure 2: CDA generated crater count statistics, showing crater frequency versus diameter.

The complete workflow for the training and application of the CDA, including generation of the characteristic surface aging curve is depicted in Figure 3.

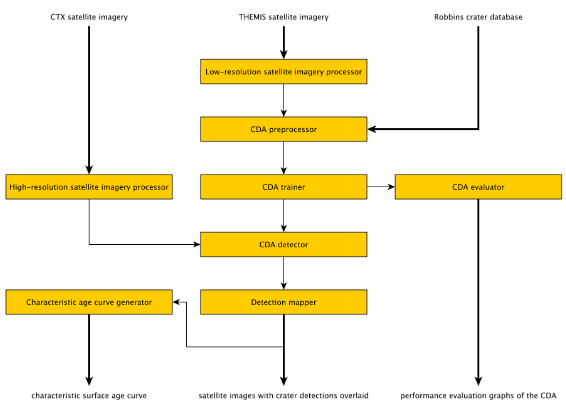


Figure 3: CDA system workflow

Conclusions: A CDA was developed, based on a Convolutional Neural Network trained using THEMIS images and the Robbins database for ground truth. The results are extremely promising, and an analysis of crater frequencies around the Mojave crater proved the concept of generalisation to higher resolution data and smaller craters.

The characteristic isochron generated automatically by the algorithm is consistent with curves generated by expert manual counting.

Future work will further refine the algorithm, and expand the training set to include examples from within a broader size range and higher resolution data. The algorithm will be applied to more surface targets, with the ultimate objective of deploying supercomputing resources to obtain crater count ages for the entire Martian surface.

References:

[1] Hartmann, W. K., and Neukum, G. (2001). *Space Science Reviews*, 96(1-4), 165-194. doi: 10.1023/A:1011945222010

[2] Kim, J. R., Muller, J. P., van Gasselt, S., Morley, J. G., and Neukum, G. (2005). *Photogrammetric Engineering & Remote Sensing*, 71(10), 1205-1217.

[3] Galloway, M. J., Benedix, G. K., Bland, P. A., Paxman, J., Towner, M. C., and Tan, T. (2014). *International Conference on Image Processing (ICIP)*, pp. 1579-1583. IEEE.

[4] Robbins, S. J. and Hynek, B. M. (2012) *Journal of Geophysical Research E: Planets*, 117(5), pp. 1–18. doi: 10.1029/2011JE003966.

[5] Robbins, S. J. et al. (2014) *Icarus*, 234, pp. 109–131. doi: 10.1016/j.icarus.2014.02.022.

[6] Stewart, R. and Andriluka, M. (2015) *Arxiv*, p. 9. doi: 10.1109/CVPR.2016.255.

[7] Werner, S. C., Ody, A. and Poulet, F. (2014) *Science*. 343(6177), pp. 1343–1346.

Interactive Gaussian Graphical Models for Discovering Depth Trends in ChemCam Data. D. A. Oyen[1], C. Komurlu[1,2], and N. L. Lanza[1]. [1]Los Alamos National Laboratory, doyen@lanl.gov. [2]Illinois Institute of Technology.

Introduction: Machine learning algorithms are now demonstrating human-level performance on some difficult benchmark problems, such as identifying objects in images. With the quantity of planetary data rapidly increasing, we would like to harness the power of machine learning to further advance planetary science. However, most of the recent successes in machine learning depend on using massive, labeled sets of data to train the algorithms. For many planetary science questions, such labeled data sets may not exist or it may not be possible to label data in such a way that would help to answer open-ended scientific inquiries. However, machine learning can improve the science return of remote sensors by increasing the speed at which scientists discover interesting patterns in their data. Interactive machine learning balances the strengths of machine learning to perform repetitive pattern recognition tasks, while empowering scientists to explore the factors that produce interesting patterns in large sets of data [1].

Spectrometers are increasingly used in remote sensing, yet spectral data can be difficult to analyze due to its high-dimensionality and non-linear mapping to interpretable quantities. As part of the Mars Science Laboratory rover operations, ChemCam's Laser-Induced Breakdown Spectroscopy (LIBS) instrument collects fine-scale atomic spectra from targets up to 7m away [2]. Given the high number of ChemCam observations to date (>400,000) and the high dimensionality of LIBS spectra (~6000 channels), advanced analysis methods are needed. We present an unsupervised machine learning method for discovering surface compositional features on rocks in ChemCam targets [3]. Our approach uses interactive machine learning to (1) give a visualization of shot-to-shot relationships among LIBS observations on a single target, and (2) identify the wavelengths (or elements) involved in the trend. Using the insight that the precision of element abundance is more reliable than accuracy [4], we bypass the quantification of elements, and look directly for patterns of chemical gradients [5, 6]. Additional trends involving different chemistry can then be explored on the same target. We are working to extend this to search the full archive of ChemCam spectroscopy data to find similar geochemical trends among all targets.

Machine Learning for Pattern Discovery: When machine learning is used for pattern discovery, we have data X but do not have labels. Therefore, we use *unsupervised* machine learning which takes the form of probability density estimation, $X \sim P(\theta)$, for a fixed distribution family P and learned parameters θ. The algorithm learns the distribution by inferring the optimal parameters $\hat{\theta}$ from the data, $\hat{\theta} = \arg\max_{\theta} \; [L(X;\theta) - R(\theta)]$, where $L(X;P(\theta))$ is the likelihood of the observed data given the probability distribution $P(\theta)$, and $R(\theta)$ is a regularization term that typically penalizes complex models. The structure of the probability distribution is typically the most interesting aspect because it reveals interesting patterns about the data. Some examples include clustering which assumes that P is a distribution with multiple modes (or centers of clusters); and probabilistic graphical models which assume that P is a multivariate joint distribution that can be factored compactly indicating direct dependencies.

Gaussian Graphical Models: Our previous work demonstrates our method for visualizing shot-to-shot relationships among LIBS observations to discover geochemical trends [5,6]. Here we give some background information to understand the approach before discussing our recent work in identifying the geochemistry involved in discovered trends. Probabilistic graphical models [7], and specifically, Gaussian graphical models (GGM) [8], are unsupervised learning models that assume that each data sample $X = (x_1, x_2, \dots, x_p)$ is a p-dimensional vector generated by a multivariate joint distribution. Furthermore, the probability distribution can be factored into a compact representation with just a few direct dependencies. The compact representation assumption is a statistical necessity for the robust estimation of a high-dimensional distribution from finite data; and it reveals interesting structure about the dependencies among variables.

To analyze the depth trend of a rock target at a location, we estimate *partial correlations* among spectra using the GGM algorithm. A partial correlation between shot A and shot B is the residual correlation after accounting for all other shots. Thus, a partial correlation is an estimate of a direct dependency. If the partial correlation between A and B is 0 then A and B are conditionally independent. A GGM is estimated from a data matrix $\boldsymbol{X}$, where each column X_j is a shot j with spectral values X_{ij} for i in $\{1, \dots, n\}$ wavelengths. The sample covariance matrix, $\boldsymbol{\Sigma}$, is calculated from $\boldsymbol{X}$, then the best sparse approximation, $\boldsymbol{\Theta}$, to the partial correlation matrix for a given sparsity constraint, λ, is estimated. The number of non-zero partial correlations is controlled by the value of λ, which can be any non-negative real number.

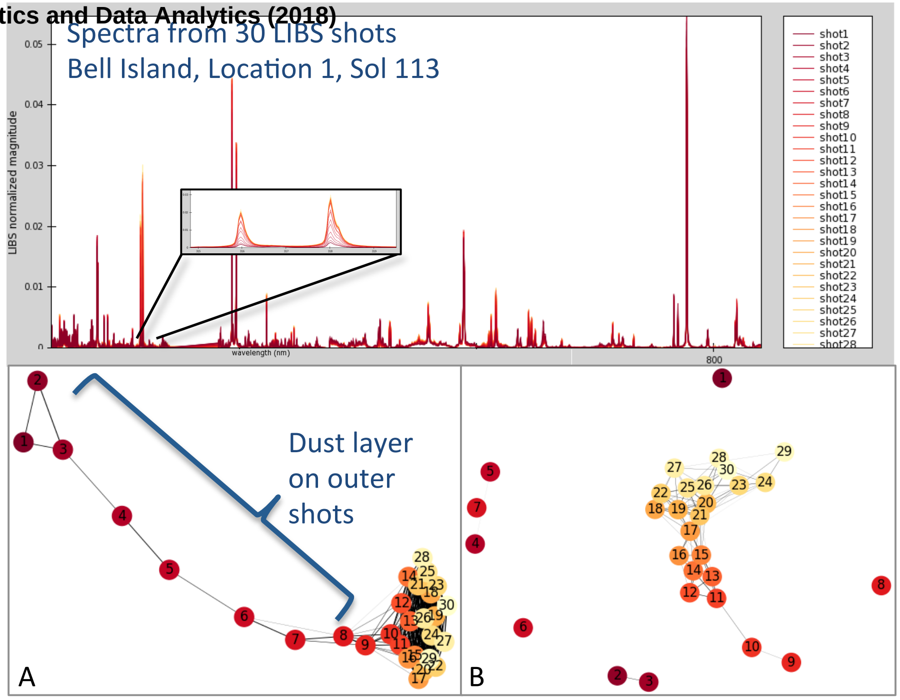


Figure 1 Interactive machine learning takes spectral data from a several LIBS shots (**top**) and learns a GGM (**bottom A**) indicating geochemical trends in ChemCam targets. In this case a clear trend is present in the first several shots at the surface of the target. iGGM then identifies the wavelengths that are responsible for the major structures of the GGM. In this case, when those wavelengths are masked, iGGM learns a GGM without that surface feature (**bottom B**), which can now be used to investigate other geochemical trends in the target.

The resulting GGM is displayed using a spring layout that places strongly correlated nodes near each other as if the correlation weights are springs pulling nodes together in space. If there are no systematic trends, then the non-zero partial correlations will appear on seemingly random pairs of shots, and the displayed GGM will look like an amorphous *blob* (or *hairball* in graph theory terminology). More visually interesting patterns emerge when there are interesting depth trends, such as a chain for systematic decrease/increase in elements, or clusters for sudden change in chemistry (such as a layer). This automated method identifies compositional depth trends associated with varnish and weathering rinds on laboratory samples [5]; and dust layers and thin sulfate veins on Mars targets [6]. We can see in the GGM Figure 1 (A) that there is a surface layer on the ChemCam target. In this case, it is a dust layer which we verified by looking at the decrease in abundance of elements associated with martian dust (e.g., Mg [10]) and the increase in abundance with S and Ca.

Explaining the Geochemistry: The GGM gives a quick visual summary of geochemical trends, but to answer specific science questions, we need to know the geochemistry behind the observed trend. We introduce an interactive Gaussian graphical model (iGGM) algorithm in which the algorithm identifies the wavelengths in the LIBS spectra that are essential for producing the most prominent structures in the learned GGM. This set of wavelengths explains the geochemistry behind the trend. For example, in the Bell Island data, the most prominent structure in the learned GGM is the chain among the first several shots. Our iGGM algorithm identifies the wavelengths that if they were masked from analysis, would make that chain disappear as in Figure 1 (B).

The iGGM algorithm identifies the critical wavelengths by searching though all possible subsets of wavelengths to find a subset of wavelengths that if they were masked would most change the structure of the learned GGM. The gradient of the weighted covariance matrix is calculated with respect to the sample weights. Then a regularization term is placed on the number of weights that can be changed to avoid the trivial solution of masking all weights. The resulting masked wavelengths are those that are critical for producing the trends seen in the GGM. iGGM also displays the newly learned GGM from the masked data as can be seen in Figure 1 which often reveals further geochemical trends in the same target.

Future Work: We plan to extend this work to facilitate quickly searching through the entire data archive of ChemCam observations to find targets with similar geochemical trends.

References: [1] Porter et al. (2013) *Comp. in Sci. & Eng.* [2] Wiens et al. (2012) *Space Sci. Rev.,* 170. [3] Lanza et al. (2015). *Icarus*. [4] Blaney et al. (2014), *JGR*, 119, 2109-2131. [5] Oyen and Lanza. (2015). *LPSC* abstract 2940 [6] Oyen and Lanza. (2017). *LPSC* abstract 1479. [7] Koller and Friedman. (2009). *Probabilistic Graphical Models*. [8] Zhao T. et al (2012) *J. Machine Learning Research*. [9] Oyen et al (2016) *Intl. Conf. Artificial Intelligence*. [10] Lasue et al. (2014). *LPSC*, abstract 1224.

PDS4: Harnessing the Power of Generate and Apache Velocity. J. Padams[1], M. Cayanan[1], S. Hardman[1], [1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA. (Jordan.H.Padams@jpl.nasa.gov)

Introduction: The PDS4 Generate Tool [1] is a Java-based command-line tool developed by the Cartography and Imaging Sciences Nodes (PDSIMG) [2] for generating PDS4 XML labels, from Apache Velocity templates and input metadata.

PDS4 is the new standard for Planetary Data System data archiving based in Extensible Markup Language (XML). This standard is used as the standard syntax and structure for capturing metadata for archived data products [3].

The metadata that data providers want to include in these metadata labels can come from many sources in many different formats: Object Description Language (ODL) detached labels, ODL attached labels, comma-separated value (CSV) files, and MySQL databases, among others.

In the past, data providers were expected to develop their own tools to parse this metadata and output these metadata labels, often duplicating effort for software that does very similar things.

The PDS4 Generate Tool provides a single package for generating PDS4 XML labels from disparate metadata sources and file formats, including ODL, CSV, and MySQL databases. In short, the Generate Tool software provides readers for the various metadata formats, and harnesses the power Apache Velocity Templating Engine [4] to output the XML label.

Apache Velocity: Apache Velocity is an open source, Java-based template engine. It permits anyone to use a simple yet powerful template language to reference objects defined in Java code [4].

For example, to produce an PDS4 XML attribute from a PDS3 Keyword, a simple variable mapping in the Velocity template can be used in order to produce the expected XML output. See Figure 1 and Figure 2 for examples of how this would occur.

Beyond the simple variable mapping, Apache Velocity also provides a user with an entire library of capabilities through the Velocity Template Language [5].

Open Source: As an open-sourced, Java-based software tool (https://github.com/nasa-pds), the PDS4 Generate Tool can be easily extended with more readers for additional metadata file formats and sources.

Consistency: With past PDS standards, data providers were left to develop their own disparate tools to generate labels for the PDS archives. In many cases this was not only duplicating effort but duplicating problems that were solved by previous data providers. This problem also left data entering the archive be inconsistent across data volumes. By using one consistent tool for this label generation, it allows data providers to help one another by growing these pipelines an integrating with other tools and services, such as the PDS Label Assistance for Interactive Design (PLAID) [6] and the PDS Label Making Tool [7].

Conclusion: This talk will provide details on the power of the PDS4 Generate Tool and Apache Velocity and how the use of this tool can benefit future data providers.

References: [1] PDS: Planetary Data System. Generate Tool. https://pds.nasa.gov/pds4/software/generate/. [2] Cartography and Imaging Sciences Node. https://pds-imaging.jpl.nasa.gov/. [3] PDS: Planetary Data System. What is PDS4? https://pds.jpl.nasa.gov/pds4/about/what.shtml [4] The Apache Velocity Project. http://velocity.apache.org/ [5] The Apache Velocity Project: User Guide. http://velocity.apache.org/engine/2.0/user-guide.html. [6] PDS Label Assistant for Interactive Design (PLAID). https://plaid.jpl.nasa.gov/ [13] De Cesare, C. et al., 2018, this volume.

NASA PDS IMG: Accessing Your Planetary Image Data.

J. Padams[1], K. Grimes[1], G. Hollins[1], S. Lavoie[1], A. Stanboli[1], and K. Wagstaff[1]. [1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA. (Jordan.H.Padams@jpl.nasa.gov)

Introduction: The Cartography and Imaging Sciences Node (IMG) of the NASA Planetary Data System (PDS) is the home to over 700 TB of digital image archives, making it one of the richest data repositories for planetary imagery in the world. Within these archives the data comes in many varieties, whether it's orbital versus landed missions, original raw experiment data versus derived products, differing coordinate systems, etc. Tools and services are needed to integrate these data so information can be correlated across missions, instruments, and data sets.

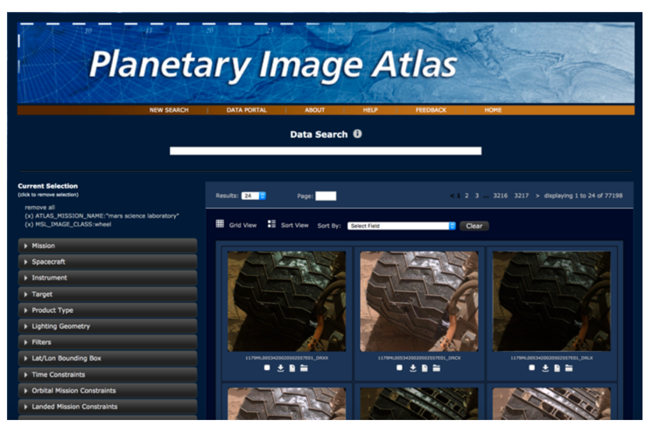
IMG has developed several tools and services to support both the wide variety of available data and the ease of access to the data both interactively, through a web browser, and programmatically through web services. From the scientist analyzing a particular crater on Mars by zooming in on a HiRISE image, to the software developer trying to build a tool to parse the metadata for all Mars Science Laboratory Hazard Camera images. The wide range of use cases provides us with a unique problem of providing interface usability for searching data, but also transparency into our backend service for software access. Leveraging partnerships with the Multimission Ground System and Service (MGSS) Office, Machine Learning and Instrument Autonomy Group (MLIA), and Multimission Image Processing Lab (MIPL) at the Jet Propulsion Laboratory (JPL) and the the expertise in planetary science, cartography, geodesy, photogrammetry and science software development at USGS Astrogeology Science Center, IMG continues to push towards new tools and services that bring the data to the people and support significant scientific discovery. For example, data archived and supported by IMG have been used to discover water on the "bone dry" Moon (Moon Mineralogy Mapper data; [1]), recent geologic activity related to CO_2 frost in martian gullies (High Resolution Imaging Science Experiment data or HiRISE; [2]), recent impacts on the Moon and Mars (Lunar Reconnaisance Orbiter Cameras or LROC; [3]; HiRISE; [4]), and recent lunar volcanism (LROC; [5]).

Atlas Web Interface: The Planetary Image Atlas (https://pds-imaging.jpl.nasa.gov/search/) provides access to the entire collection of IMG data through links to online holdings and data node catalogs [6]. The PDS Imaging Node Atlas utilizes faceted navigation, an interactive style of browsing datasets that allows users to filter a set of items by progressively selecting from only valid values of a faceted classification system. In the Atlas, facets are defined by the most commonly used search criteria for imaging datasets including but not limited to: mission name, instrument name, target, product type, lighting geometry meta-data (emission angle, incidence angle, phase angle), lat/lon meta-data, time constraints, etc. In addition to the faceted approach, the Atlas builds on the features of the previous Atlas including a map interface for the Saturnian moons, Earth's moon and Mars. The Atlas also incorporates the use of the MGSS webification backend that makes use of the image transformation software developed by MGSS (MIPL) through JavaScript widgets [7].

A recent, powerful enhancement to the Atlas is the ability to search for images based on the "content" in the image (i.e. crater, moon, rings). In order to detect content in the images, we developed software that leverages deep convolutional neural networks (CNN). "CNNs organize image content in increasingly complex representations starting at the pixel level up to entire objects, such as rings, craters, moons, and so on. Lower-level information, such as edges, corners, etc., are common to all content. The specifics of how the low-level information gets combined into high level representations is unique to the domain and requires training of the network with target content and associated labels." [8] Users can leverage this image content search to investigate particular craters on mars or study the wheel wear on the Mars Science Laboratory Curiosity rover [9].

Another recent enhancement of the Atlas is an improved Google-like search thats allows users the flexibility of simple search syntax with a refined result set, like 'mars crater' or 'cassini moon'. Users can then further narrow their search using the faceting features or additional text searches.

All the search capabilities of the Atlas, both new and old, help users across the globe find and download nearly 15 TB of data each month.



Atlas Search Service: The Atlas Web Interface displays query results returned from the Atlas Search Service (https://pds-imaging.jpl.nasa.gov/solr/pds_archives/search), a web service extending Apache Solr [10]. Harnessing the power of Solr, this web service provides fast search, JavaScript Object Notation (JSON) and eXtensible Markup Language (XML) return types for software parsing, and a refined query syntax for complicated searches. In addition, the PDS IMG Search Protocol was developed to provide simplified query access to the data through a REST-like API, as well as provide a means for integrated search across the PDS. As an extension of the PDS Engineering Node (PDS EN) Search Protocol [11], it heavily leverages the query parser syntax from Apache Lucene [12] as well as certain characteristics from the Planetary Data Access Protocol (PDAP) [11] developed by the International Planetary Data Alliance (IPDA). The protocol is intended for finding large sets of data, for example, searching by instrument, instrument-host, instrument-type, start-time, stop-time, etc. The PDS IMG Search Protocol (https://pds-imaging.jpl.nasa.gov/search/protocol) extends this by providing search against more imaging product-specific metadata, such as product-type, filter, image-content, orbit, planet-day, etc.

Webification (w10n): Webification (W10N) (https://pds-imaging.jpl.nasa.gov/tools/w10n/) is a specification that defines a common way to expose resources (composite files, databases, command-line applications, etc.) on the web.

The core idea is to make the inner components of resources directly addressable and accessible via well-defined and semantically meaningful URLs. The MGSS Web Resources Platform (WRP) provides a set of tools that leverage this specification to provide access to data and applications (services) through ReSTful URLs [13]. PDS IMG uses the Juneberry component of WRP to allow simple programmatic access to the data archive (https://pds-imaging.jpl.nasa.gov/w10n/), and Servicification (Serv10n) for access to backend services. This service is central to the server-side functionality for several IMG services, including the Planetary Image Atlas, PDS Marsviewer, and Landmarks Web Services.

References: [1] Pieters, C.M. et al., 2009, Science, v. 326, #5952, pp. 568-572. [2] Dundas, C.M. et al., 2012, Icarus 220, pp. 124-143. [3] Robinson, M.S. et al., 2015, Icarus 252, pp. 229-235. [4] Dundas, C.M. et al., 2014, JGR-P, 119, 109-127. [5] Braden, S. et al., 2014, Nature Geoscience, v. 7, 787-791. [6] Gaddis, L., et al., 2014, USGS Open-File Report 2014-1056, p. 197-199. [7] Stanboli, A. et al., 2015, Proceedings of the 2nd Planetary Data Workshop. [8] Altinok A. et al., 2017, Proceedings of the 3rd Planetary Data Workshop. [9] http://www.planetary.org/blogs/emily-lakdawalla/2014/08190630-curiosity-wheel-damage.html. [10] http://lucene.apache.org/solr/. [11] Hardman, S., et al., 2015, Proceedings of the 2nd Planetary Data Workshop. [12] https://lucene.apache.org/core/2_9_4/queryparsersyntax.html. [13] Grimes, K. et al., 2017, Proceedings of the 3rd Planetary Data Workshop.

Example Atlas Search Service URLs
Query for images of Mars
https://pds-imaging.jpl.nasa.gov/solr/pds_archives/search?target=mars
Query for images of Mars taken by Curiosity instruments
https://pds-imaging.jpl.nasa.gov/solr/pds_archives/search?target=mars&spacecraft=curiosity
Query for mosaics of Mars taken by Curiosity instruments
https://pds-imaging.jpl.nasa.gov/solr/pds_archives/search?target=mars&spacecraft=curiosity&product-type=mosaic

Example W10n URLs
To display metadata info, i.e., content of Cassini ISS data directory as JSON output
https://pds-imaging.jpl.nasa.gov/w10n/cassini/cassini_orbiter/coiss_2015/data/?output=json
Return meta info of image 0 in JSON. One attribute metadata contains VICAR labels
https://.../coiss_2015/data/1506288646_1506388236/N1506378403_1.IMG/0/
Returns image 0 in GIF format.
https://.../coiss_2015/data/1506288646_1506388236/N1506378403_1.IMG/0/image[]?output=gif
Return image 0 raster data array in JSON. Or in big/little-endian binary if output is set to big-endian or little-endian.
https://.../coiss_2015/data/1506288646_1506388236/N1506378403_1.IMG/0/raster/data[]?output=json

Ground observation of asteroids at mission ETA. F. Paganelli[1] and A. Conrad[2], [1] SETI/NCCC (fpaganelli@nccc.edu), [2]LBTO/UHH (aconrad@lbto.org).

Introduction: NASA funding for ground-based support of spacecraft missions to small bodies has increased to enhance data science return.Thus emphasis for targeted asteroids (i.e. Lucy) data collected during observations carried out at ground-based observatories at mission estimated time of arrival (ETA), would provide to, and be recognized by, NASA as a valuable asset. Lucy, a SwRI mission proposal to study primitive asteroids among the Jupiter's Trojans, is one of five science investigations under the NASA Discovery Program [1]. Lucy's science payload/instrumentation is mirroring the New Horizons payload with: L'Ralph, Panchromatic and color visible imager and Infrared spectroscopic mapper (400nm -2.5µm); L'LORRI, high-resolution visible imager (350-850 nm); L'TES, thermal infrared spectrometer is similar to OTES on the OSIRIS-REx mission (spectral range 5.71–100 µm (1750–100 cm–1); and radio science investigation will determine the mass of the Trojans by using the spacecraft radio telecommunications hardware to measure Doppler shifts [2].

Approach: Observations through the LBT Multi-Object Double Spectrographs (MODS 1) - Imager and spectrograph covering 0.32-1.1 microns with a 6'x6' FOV - has been targeted for this assessment [3]. The importance of some Eurybates members spectra show a drop off in reflectance shortward of 0.52µm - similar features are seen in main belt C-type asteroids and commonly attributed to the intervalence charge transfer transition in oxidized iron [4,5].

Mission	Number	Target	Vmag	Size (mas)	Albedo (µm)
Lucy	3548	Eurybates	16.8 to 17.7	13 to 20	0.052
	15094	Polymele	18.9 to 19.8	5 to 7	0.091
	11351	Leucus	17.8 to 18.8	7 to 11	0.079
	21900	Orus	16.9 to 17.9	11 to 16	0.075
	617	Patroclus	15.9 to 16.5	33 to 39	0.047
	52246	Donaldjohanson	18.3 to 20.1	2 to 4	

Figure 1. Lucy's Jupiter Trojan Asteroids albedo [4,5].

To derive Lucy targeted asteroids information for best ground-based observation at mission estimated time of arrival (ETA) we used data from JPL Horizons [6].

Lucy Mission	Encounter date	Location	Dia-meter (km)	Spectral type	**ETA targets**
Launch	Oct. 2021				
DonaldJohanson	April 2025	Main belt	3.9	C	
Eurybates	Aug. 2027	Greeks	64	C	12 Aug 2027
Polymele	Sept. 2027	Greeks	21	P	15 Sep 2027
Leucus	April 2028	Greeks	34	D	18 Apr 2028
Orus	Nov. 2028	Greeks	51	D	11 Nov 2028
Patroclus/Menoetius	March 2033	Trojans	113/ 104	P	02 Mar 2033

Figure 2. Lucy's Jupiter Trojan Asteroids ETA [1,2].

The workflow, shown in Figure 3, used *expect & tcl*, plus a *python* wrapper to access the JPL Horizons [6] database and extract observations of targeted asteroids at twilight conditions. The derived data provide the best suitable opportunities to observe the asteroids using LBT ground observations.

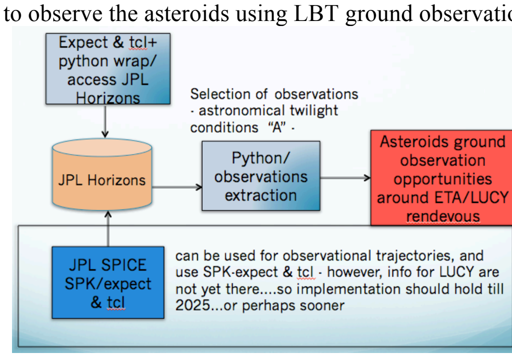


Figure 3. Workflow/pipeline for data extraction.

Results: The extracted observations for all targeted asteroids outlined several opportunities for suitable LBT ground observations. However, only one was found to be suitable during close approach of Lucy ETA to asteroid Leucus, as shown in Figure 4.

```
Leucus
 Date___UT___HR:MM:SS Date____JDUT       A/  R.A.___(ICRF/J2000.0)___DEC dRA*cosD  d(DEC)/dt  APmag S-brt Azi_(a-appr)_Elev
 2028-Apr-01 12:00:01 2461863.000011574 A   16 48 05.3186 -23 52 34.935 -2.03708  1.967876  18.61   8.16 188.7451  32.8560
 2028-Apr-02 12:00:01 2461864.000011574 A   16 48 02.0368 -23 51 47.257 -2.47764  2.039883  18.60   8.16 189.8220  32.7334
 2028-Apr-03 12:00:01 2461865.000011574 A   16 47 57.9864 -23 50 57.883 -2.91644  2.112077  18.59   8.15 190.8982  32.5952
 2028-Apr-04 12:00:01 2461866.000011574 A   16 47 53.1706 -23 50 06.809 -3.35331  2.184463  18.58   8.15 191.9734  32.4413
 2028-Apr-05 12:00:01 2461867.000011574 A   16 47 47.5929 -23 49 14.030 -3.78812  2.257043  18.57   8.15 193.0469  32.2716
 2028-Apr-06 12:00:01 2461868.000011574 A   16 47 41.2569 -23 48 19.543 -4.22079  2.329823  18.56   8.14 194.1184  32.0863
 2028-Apr-07 12:00:01 2461869.000011574 A   16 47 34.1662 -23 47 23.343 -4.65131  2.402808  18.55   8.14 195.1872  31.8853
 2028-Apr-11 03:00:01 2461872.625011574 A   16 47 02.3258 -23 43 44.342 -5.64235  2.480793  18.51   8.12  97.9176 -34.7549
 2028-Apr-12 03:00:01 2461873.625011574 A   16 46 51.7975 -23 42 40.171 -6.06957  2.550339  18.50   8.12  98.3705 -33.8900
 2028-Apr-13 03:00:01 2461874.625011574 A   16 46 40.5368 -23 41 34.262 -6.49439  2.620069  18.49   8.11  98.8227 -33.0234
 2028-Apr-14 03:00:01 2461875.625011574 A   16 46 28.5406 -23 40 26.614 -6.91652  2.689957  18.48   8.11  99.2744 -32.1550
 2028-Apr-15 03:00:01 2461876.625011574 A   16 46 15.8382 -23 39 17.225 -7.33564  2.759975  18.47   8.10  99.7260 -31.2850
 2028-Apr-16 03:00:01 2461877.625011574 A   16 46 02.4117 -23 38 06.092 -7.75140  2.830090  18.46   8.10 100.1776 -30.4134
 2028-Apr-17 03:00:01 2461878.625011574 A   16 45 48.2757 -23 36 53.215 -8.16346  2.900269  18.44   8.09 100.6294 -29.5403
 2028-Apr-18 03:00:01 2461879.625011574 A   16 45 33.4374 -23 35 38.595 -8.57145  2.970479  18.43   8.08 101.0817 -28.6658
 2028-Apr-19 03:00:01 2461880.625011574 A   16 45 17.9047 -23 34 22.233 -8.97501  3.040684  18.42   8.08 101.5347 -27.7899
 2028-Apr-20 03:00:01 2461881.625011574 A   16 45 01.6862 -23 33 04.131 -9.37380  3.110844  18.41   8.07 101.9885 -26.9129
 2028-Apr-21 03:00:01 2461882.625011574 A   16 44 44.7908 -23 31 44.293 -9.76743  3.180921  18.40   8.07 102.4433 -26.0347
 2028-Apr-22 03:00:01 2461883.625011574 A   16 44 27.2284 -23 30 22.723 -10.1555  3.250870  18.39   8.06 102.8995 -25.1554
 2028-Apr-23 03:00:01 2461884.625011574 A   16 44 09.0094 -23 28 59.426 -10.5377  3.320647  18.38   8.05 103.3570 -24.2752
 2028-Apr-24 03:00:01 2461885.625011574 A   16 43 50.1450 -23 27 34.408 -10.9136  3.390205  18.37   8.05 103.8162 -23.3941
 2028-Apr-25 03:00:01 2461886.625011574 A   16 43 30.6468 -23 26 07.678 -11.2828  3.459496  18.35   8.04 104.2773 -22.5123
Observation days for Leucus are 22
```

Figure 4. LBT ground observations.

Considerations: This effort could be enhanced by integrated spacecraft, space telescope, and ground observatory missions. A possible space telescope would be SOFIA, while candidate ground station to be considered in future work is the European Southern Observatory (ESO) in Chile. Also, the 23m Fizeau Imaging on LBT could fill the pre-ELT gap (~ 2018 to 2023) for resolved imaging of Lucy mission targets via appulse events, which are estimated to occur approximately once per week [7].

References: [1] Levison H.F. and the Lucy Science team (2016) *LPSC 47th,* Abstract #2061. [2] Weaver H.A. et al. (2008) *Space Sci. Rev. 140(1-4),* 75-91. [3] Rothberg B. et al. (2016) Astrophysics, arXiv:1608.00037 [astro-ph.IM]. [4] Fornasier et al. (2007) *Icarus, 190 (2),* 622-642. [5] Fernadez Y.R. et al. (2009) *The Astro. J.* 138, 240-250. J. H. (1996) *LPS XXVII,* 1344–1345. [6] JPL Horizons: ftp://ssd.jpl.nasa.gov/pub/ssd/Horizons_doc.pdf
[7] Conrad A. et al. (2017) *AO4ELT5,* 1-8.

RAMAN/LIBS DATA FUSION VIA TWO-WAY VARIATIONAL AUTOENCODERS Mario Parente[1], I. Gemp[2], [1]ECE Department, University of Massachusetts, 01003 Amherst, MA; mparente@ecs.umass.edu; [2] CICS, University of Massachusetts, 01003 Amherst, MA

Introduction: Several upcoming missions will feature Raman and LIBS spectroscopy for planetary exploration on Mars (e.g. Mars 2020 SuperCam and SHERLOC). Supporting the scientific effort will require a deeper understanding how to interpret Raman spectra of mineral assemblages in soils and rocks. Unfortunately, there is no underlying theoretical understanding of mixtures for Raman spectroscopy, and developing the necessary theoretical basis for understanding Raman mixing phenomena will take decades.

As an alternative, we are investigating a completely different, original solution to extracting mineral abundances from Raman data by combining Raman results with LIBS using deep learning and data fusion. Raman alone constrains what phase is present and the relative cation proportions in each phase. LIBS provides the total numbers of those cations, so the combination of the two allows the proportion of each phase (the modal abundance) to be obtained. This requires only knowledge of the mineral formulas of each phase identified by Raman, the associated algorithm for interpreting its individual composition from peak position and the LIBS total amounts of those cations.

***Two-way* Variational Autoencoder [1, 2, 3]:** We consider the marriage of two probabilistic models to describe the data. The first is a probabilistic model (M2 in [4]) that describes the spectra, $\mathbf{x}$, as being generated by a composition vector[1] $\mathbf{y}$ in addition to a latent, nuisance vector $\mathbf{z}$. The joint distribution is assumed to factorize as $p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = p(\mathbf{y})p(\mathbf{z})p(\mathbf{x}|\mathbf{y}, \mathbf{z})$, so the data are explained by the *generative process*:

$$p(\mathbf{y}) = Dir(\mathbf{1}) \tag{1}$$

$$p(\mathbf{z}) = U(-\mathbf{1.5}, \mathbf{1.5}) \tag{2}$$

$$p_\theta(\mathbf{x}|\mathbf{y}, \mathbf{z}) = f(\mathbf{x}; \mathbf{y}, \mathbf{z}, \theta) \tag{3}$$

Here, $p(\mathbf{y})$ and $p(\mathbf{z})$ are prior distributions and $f(\mathbf{x}; \mathbf{y}, \mathbf{z}, \theta)$ is a distribution whose parameters are non-linear functions of $\mathbf{y}$ and $\mathbf{z}$ (e.g., diagonal Gaussian $\mathcal{N}(\mu_\theta(\mathbf{y}, \mathbf{z}); \mathbf{\Sigma}_\theta(\mathbf{y}, \mathbf{z})))$. We choose a uniform prior over the simplex for compositions, $Dir(\mathbf{1})$, and deep neural networks with weights θ for $\mu_\theta(\mathbf{y}, \mathbf{z})$ and $\mathbf{\Sigma}_\theta(\mathbf{y}, \mathbf{z})$. The second is a probabilistic model that describes the *reverse process*: nuisances and compositions are generated by spectra,

$$q(\mathbf{x}) = U(-\gamma, \gamma) \text{ e.g., } \gamma \gg 0 \tag{4}$$

$$q_\phi(\mathbf{y}|\mathbf{x}) = g(\mathbf{y}; \mathbf{x}, \phi) \tag{5}$$

$$q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y}) = h(\mathbf{z}; \mathbf{x}, \mathbf{y}, \phi) \tag{6}$$

where $q(\mathbf{x})$ is an uninformative, uniform prior, $q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{y})$ is a diagonal Gaussian parametrized by a deep neural network as before, and $\gamma \gg 0$. To define $q_\phi(\mathbf{y}|\mathbf{x})$ with support limited to the simplex, we first draw an intermediate random variable, $\tilde{\mathbf{y}}$, from a diagonal Gaussian. We then pass $\tilde{\mathbf{y}}$ through a non-linear transformation called a *normalizing flow* [5] which conforms the Gaussian distribution to the simplex.

To learn the parameters, θ and ϕ, we optimize variational lower bounds on the marginal likelihoods of our data samples [1, 2, 3]. The marginal likelihood for the entire dataset is

$$\mathcal{J}_f = \sum_{(\mathbf{x},\mathbf{y})\sim\tilde{p}_l} \mathbb{L}_f(x, y) + \sum_{\mathbf{x}\sim\tilde{p}_{u_x}} \mathbb{L}_f(x) \tag{7}$$

$$\mathcal{J}_r = \sum_{(\mathbf{x},\mathbf{y})\sim\tilde{p}_l} \mathbb{L}_r(x, y) + \sum_{\mathbf{y}\sim\tilde{p}_{u_y}} \mathbb{L}_r(x) \tag{8}$$

for the forward and reverse models, respectively [1, 2, 3]. As in [4], we introduce an additional discriminative objective to each model that can be learned from the labeled data:

$$\mathcal{J}_f^d = \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim\tilde{p}_l} L(\bar{\mathbf{y}}, \mathbf{y}), \quad \mathcal{J}_r^d = \mathbb{E}_{(\mathbf{x},\mathbf{y})\sim\tilde{p}_l} L(\bar{\mathbf{x}}, \mathbf{x})$$

where $\bar{\mathbf{y}}$ can, for example, either be a sample from $q_\phi(\mathbf{y}|\mathbf{x})$ or the mean of the distribution and L can, for example, be KL($\mathbf{y} \parallel \bar{\mathbf{y}}$); $\mathbf{x}$ is treated similarly using $p_\theta(\mathbf{x}|\mathbf{y}, \mathbf{z})$ and $L = \parallel \bar{\mathbf{x}} - \mathbf{x} \parallel^2$. We jointly optimize both models as a weighted sum:

$$\mathcal{J} = \alpha_f \mathcal{J}_f - \alpha_f^d \mathcal{J}_{\mathrm{f}}^{\mathrm{d}} + \alpha_r \mathcal{J}_r - \alpha_r^d \mathcal{J}_{\mathrm{r}}^{\mathrm{d}}. \tag{9}$$

We learn the parameters θ and ϕ by maximizing (9) using Monte Carlo samples for the latent variables —a technique known as stochastic gradient variational Bayes [6] or stochastic backpropagation [7].

Data Fusion: The data fusion model incorporates both LIBS and Raman data simultaneously. The approach requires that LIBS and Raman spectra be collected from the same samples, thereby creating a dataset consisting of N trials, each measuring spectral intensities at $S = L + R$ channels ($L = \#$ of LIBS channels, $R = \#$ of Raman channels) with corresponding mineral abundances for M minerals. Raman spectra are captured for 567 binary mineral mixtures created in the lab. For our training dataset, $N = 189$ samples, $M = 23$ mineral types, $L = 5485$ LIBS channels, and $R = 1715$ Raman channels. We hold out 378 samples to validate our data fusion approach. We use the

[1]By composition we will indicate mineral abundances for Raman data or elemental compositions for LIBS data.

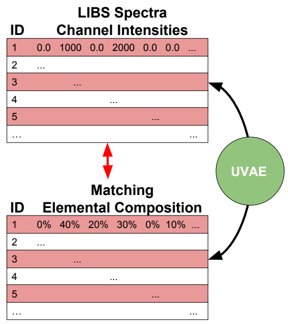


Figure 1: We have 130 *LIBS lab spectra obtained under Mars conditions with matching known elemental compositions. We also have* 500 *additional samples acquired in the same conditions with unknown elemental compositions. Each spectrum contains* 5485 *channels. By training the UVAE model to capture the relationship between these two representations, we acquire the ability to generate* synthetic *LIBS spectra. Specifically, training the UVAE results in learning the distribution over LIBS spectra given elemental composition,* $p(\mathbf{x}|\mathbf{y})$*. We can synthesize a new LIBS spectrum,* $\mathbf{x}$*, from this distribution conditioned on a specific elemental composition,* $\mathbf{y}$*.*

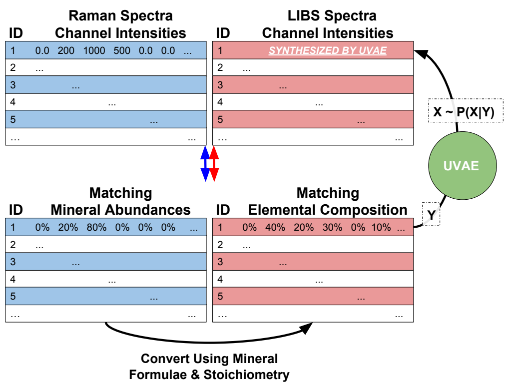


Figure 2: We synthesize LIBS spectra for our mineral abundance prediction challenge by 1) converting mineral abundances to elemental composition using stoichiometry, 2) training the UVAE to model the relationship between LIBS spectra and elemental composition, and then 3) generating LIBS spectra conditioned on these elemental compositions using the UVAE model trained in step 2.

deep generative model described above, namely the Untapped Variational Autoencoder (UVAE), to predict the percent mineral abundances for 23 minerals given the S-dimensional LIBS+Raman spectral data.

*Synthesizing an Appropriate Dataset***:** The *data fusion* scenario we consider assumes both Raman spectra and LIBS spectra have been acquired in conjunction with matching mineral abundance for each rock sample. In practical in situ analyses one might only have access to LIBS spectra and corresponding percent elemental compositions for some samples and Raman spectra corresponding mineral abundances on other samples. In order to simulate this scenario, we devise a process for synthesizing matching LIBS spectra. First, we train a UVAE to model the relationship between LIBS spectra and elemental composition (see Figure 1). Next, we use stoichiometry to convert atomic weights for each mineral to percent elemental composition. For each measurement trial in our Raman dataset, we calculate the percent elemental composition of a mineral mixture by combining the elemental compositions of each mineral in corresponding proportions. Then, using the trained UVAE model, we generate LIBS spectra conditioned on the derived elemental compositions from the Raman dataset. This results in a LIBS+Raman spectral dataset where the LIBS portion of the spectrum is synthetic, however, deliberately crafted to mimic the real matching LIBS spectra although with sometimes significant loss in fidelity. The process is shown in Figure 2.

*Predicting Mineral Abundances***:** We compare three models: 1) a UVAE trained on Raman+LIBS spectra, 2) a UVAE trained on Raman only, and 3) PLS trained on Raman only. In order to make predictions of mineral abundances using the UVAE model, we utilize $q(\mathbf{y}|\mathbf{x})$. We can either sample from this distribution or take the mean of this distribution; in our results we use the mean. Preliminary results show that including the LIBS spectral data (even though synthetic!) is helpful in driving down the error of predicting mineral abundances. Here,

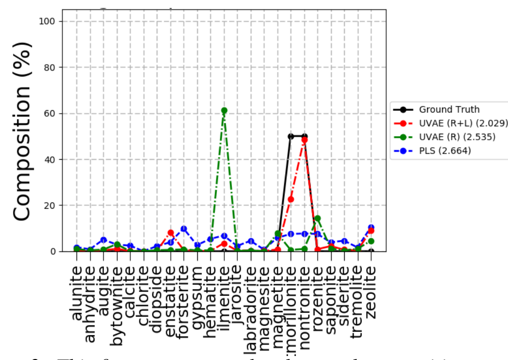


Figure 3: This figure compares the elemental composition predicted by each model for a single rock sample. The average KL divergence for each model is given beside the model name in the legend. Mineral names are on the $\mathbf{x}$*-axis while % abundance is on the* $\mathbf{y}$*-axis.*

we measure error as the average KL divergence between the true mineral abundance of a rock sample and the predicted mineral abundance given by the model (lower is better) on the held out validation set. Figure 3 shows the percent mineral abundance for a selected sample as predicted by UVAE (red), PLS (blue), and actual (black). The average KL divergence over all samples in the validation set is quantified for each model in the legend.

References: [1] I. Gemp, et al. (2016) *ArXiv e-prints*. `arXiv:1608.05983`. [2] I. Gemp, et al. (2017) in *NIPS AABI*, Long Beach, CA, USA. [3] M. Parente, et al. (2017) in *IEEE IGARSS* Forth Worth, TX, USA. [4] D. Kingma, et al. (2014) in *NIPS* 3581–3589. [5] D. J. Rezende, et al. (2015) *arXiv preprint arXiv:150505770*. [6] D. Kingma, et al. (2013) *arXiv preprint arXiv:13126114*. [7] D. J. Rezende, et al. (2014) *arXiv preprint arXiv:14014082*.

SCALABLE DATA PROCESSING WITH THE LROC PROCESSING PIPELINES. K. N. Paris, N. M. Estes, E. Cisneros, and M. S. Robinson, School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287.

Introduction: The Lunar Reconnaissance Orbiter Camera (LROC) is a suite of three cameras onboard the Lunar Reconnaissance Orbiter (LRO), which has been systematically mapping the Moon since 2009 [1]. The LROC includes two Narrow Angle Cameras (NACs) and one Wide Angle Camera (WAC) [2] that collect hundreds of images each day, totaling over 2.2 million observations as of November 2017.

In addition to the science files, the LROC Science Operations Center (SOC) generates and receives from the Mission Operations Center (MOC) almost 100 other types of files. Many of the file types are necessary to plan observations, facilitate science file processing, and generate products for the scientific community and the public. Keeping track of these diverse file types requires a well thought out and scalable process for future expansion or compression: additional file types, additional processing steps to pipelines, retiring pipelines, and future mission support.

Processing Strategy: There are four components that have made the processing strategy for the LROC SOC as successful as it is. These are 1) the processing pipelines, 2) a mission database, 3) a job and resource management software suite, and 4) a robust file system.

Pipeline Scripts aka Procedures: Individual scripts are referred to as "procedures" and collections of procedures relating to a particular file type are referred to as "pipelines". Each procedure is written to accomplish a very specific task and can chain to other procedures as necessary. This approach was taken to keep the scripts uncomplicated for the ease of development and reprocessing purposes, and to be able to simplify error handling (reduce duplicate processing or rolling back any processing performed).

The procedures are written in Bash. Bash was chosen as the preferred language for read-ability by the Operations staff (who aren't necessarily software developers), who would be creating or updating the pipelines. In some of the Bash scripts, there are calls to programs that are written in C++ or Ruby where appropriate (i.e. image processing or geometry calculations).

Whenever a file is received or generated by the LROC SOC it is "ingested" – this is the process by which the file and ancillary data are cataloged in a database. After a file is ingested, it is passed to the "pipeline" for that file type where further processing occurs via individual procedures.

The pipelines and procedures are named in a way that makes it easy to identify which file type that pipeline processes and the order in which procedures are to be executed. For example (Fig. 1): lroc_moc27_p30 is the third step (p30) in processing of the moc27 file type.

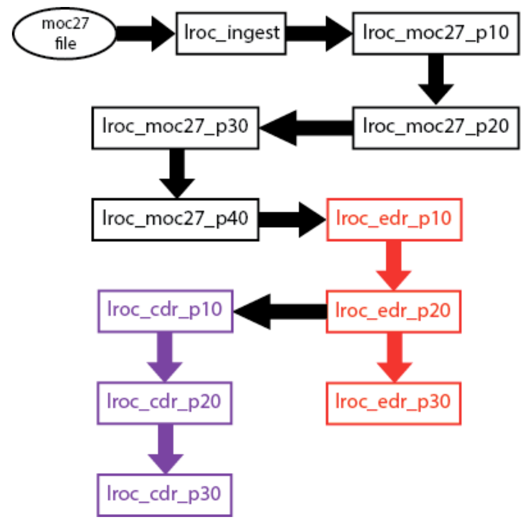


Figure 1: Example of the processing flow for a moc27 file type through the moc27 pipeline (black), EDR pipeline (red), and CDR pipeline (purple).

Figure 1 is an example of a file being processed through the several relevant pipelines. A science file is received from the MOC (file type moc27) and submitted to the ingest script where it is initially cataloged It is then passed to the first step in the moc27 processing pipeline which includes four procedures (lroc_moc27_p10 – p40).

The first procedure compares the values in the science file header to the expected values that are stored in the database, and creates and updates fields in the database to indicate that a science file for an observation has been received. The p10 procedure then submits the science file to the second (p20) procedure, which updates specific timing-related fields in the database for the observation and passes the file onto the p30 script. The third procedure (p30) checks the file for validity and computes statistics from the image data and sends the file to the final procedure in the moc27 pipeline (p40). The fourth procedure (p40) reviews the image statistics, checks for housekeeping and SPICE data and computes the data quality id for the observation. Assuming that all of the necessary data are present for the science file, the lroc_moc27_p40 procedure then submits the science file to the EDR pipeline, which starts with the lroc_edr_p10 procedure. Assuming success in the p10 and p20 procedures, the file is submitted to the final procedure in the EDR pipeline and also to the first procedure in the CDR pipeline.

Whenever a job fails, the user is made aware of it via the Rector interface [3]. The output of the script is saved for review and mitigation development. Once the necessary mitigation steps have been taken for successful processing, the job is resubmitted to the pipeline that it failed out of and processing can proceed.

Database: The LROC SOC uses a PostgreSQL database to track received files and track the status of the processing for the files. A record is created for each file received and meta-data captured from the file. Most commonly, this meta-data consists of the start and end times of a file, the delivery time, file md5 checksum and file path in the file system.

Resource Management. Unless paused, the pipelines are always "on" and available to process whenever jobs are submitted. This requires careful resource management on the processing cluster, which is done by an in-house developed program called Rector. More information about the Rector software can be found in the "Taming Pipelines, Users, and High Performance Computing with Rector" [3].

File System: The LROC SOC maintains a fully redundant 2.3 Petabyte storage array where project files are organized and stored. The storage array is shared via the NFS protocol to our SOC computer cluster, as well as mission workstations to facilitate sharing LROC files. Cluster nodes make use of locally attached disks for pipeline processing. If new file products are generated, that pipeline step will move the file product to the shared storage array.

Development and Testing: The LROC pipelines are under version control using Subversion control software [4]. Any script that is related to the pipelines and the pipelines themselves are stored in the same subversion project. When updates are made to any of the software, it is deployed to a separate test cluster, integration testing performed, and when successful, committed to the subversion project. Once testing is successful and the changes committed, a new release is tagged, the affected pipelines are paused by the Operations staff, and the newly tagged subversion project is deployed to the production file system.

Evolution of the Pipelines: The pipelines started off as a simple set of scripts for ingesting and processing files as they are received. As the mission continued and operations evolved the pipelines have expanded into a set of 132 procedures to handle new file types requested from the MOC, some error handling situations could be handled automatically, and added steps to increase the usability of the data for the greater science community. Several key lessons learned during the evolution process are described in detailed below.

Timing Challenges: Given the uncertainty in timing of file delivery and the dependencies of some file types on others, one of the biggest challenges faced was how to process files in the right order when they aren't necessarily received in an required order. For example, processing a science file requires SPICE files and housekeeping telemetry files which may come before or after the science file itself.

To mitigate the problem of file delivery timing, meta-data for each file are stored in the database, including file type for each file and the start and end times for files where those data exist. A separate non-pipeline script was written and uses this meta-data to determine if given file types exists for a particular observation or science file. This script is called by each pipeline for which there is a file processing dependency. In the example of the science file processing, the script is called in the science file pipeline, SPICE file pipelines, and in the housekeeping pipeline. This ensures that no matter what the delivery times of the files are, the science file can be processed automatically.

Keep multiple instances of a job from running: In situations when one job is resource-intensive, multiple instances that job running at the same time was be problematic. This would cause computers on the cluster or the production database to greatly slow down, which affects the Operations staff and anyone else utilizing the processing cluster. To deal with this, lock files were implemented so that when one instance of a job starts, it checks for a lock file before starting the intensive work and if a lock file exists, the job waits and checks for the lock file's existence at set intervals (i.e. 30-60 seconds). If a lock file does not exist, the script creates a lock file to prevent other instances of the job from running.

Reprocessing: Files need to be reprocessed as calibration is updated, geometry is improved, or for any other myriad of reasons. The pipelines were written in such a way that jobs can be submitted to a single procedure to kick-off reprocessing of particular files or file types.

Future work: Future work will focus on transferring knowledge and processes to new missions currently in development, as well as continued refinement for the LRO mission.

References: [1] Chin et al., Space Sci Rev (special issue. [2] Robinson et al., Space Sci Rev. [3] Estes et al., submitted this volume. [4] https://subversion.apache.org/

LARGE-SCALE NUMERICAL SIMULATIONS OF PLANETARY INTERIORS. A. C. Plesa[1], M. Maurice[1], S. Padovan[1], N. Tosi[1,2], D. Breuer[1]. [1]German Aerospace Center (DLR), [2]Technische Universität Berlin.

Introduction: The large amount of data returned from space missions and telescopes has helped to improve our understanding of the thermochemical evolution of terrestrial planets in our Solar System and beyond. However, the interior dynamics of terrestrial bodies is poorly constrained as direct constraints are lacking and laboratory experiments can only cover a limited parameter space often not representative of the deep interior. Over the past decades, large-scale numerical simulations of interior evolution have grown to become one of the most powerful approaches to model the temporal evolution of the mantle flow in the interior of terrestrial planets. Large improvements both in processor power and amount of available memory have made possible to tackle highly complex scenarios of planetary evolution by modeling vigorous convection at extreme Rayleigh numbers with highly variable viscosity and including chemical heterogeneities in a 3D spherical geometry.

Modeling the interior evolution of terrestrial planets: The thermochemical evolution of the mantle of terrestrial planets is modeled by solving the fundamental equations governing the conservation of mass, momentum and energy [1]. Convection driven by active chemical fields is modeled with the particle in cell method (PIC) [e.g., 2, 3, 4]. This method employs Lagrangian tracers that carry various chemical species, and has the advantage over classical grid-based methods of being essentially free of numerical diffusion. It also provides the advection of an arbitrary large number of different compositional fields with less computational effort.

We use the fluid flow solver Gaia, which is written in library-independent C++ (except MPI that can be disabled), making it easy to port on all type of systems. Gaia uses the finite-volume method to discretize the governing equations on arbitrary grids in various geometries as long as they are Voronoi grids [5, 6].

An efficient domain decomposition of a given computational mesh results into *n* equal volumes, with each volume being mapped to a computational core. While for 3D Cartesian box grids, due to their regular nature, an optimal domain decomposition can be easily achieved, the problem becomes highly complex when 3D spherical shells are involved. In Gaia we use the so-called Thomson-points to laterally decompose the sphere by distributing points, all assumed with equivalent potential energy, on the surface of the sphere and minimizing the global potential field energy. The closest Thomson-point defines the domain of every grid cell [7, 8].

We tested the performance of our code on the Hazel Hen system of HLRS using up to 54 x 10^3 computational cores (dual socket Intel(R) Xeon(R) CPU E5-2680 v3 @ 2.5 GHz having two sockets per node with 12 cores each). To this end we have performed numerical simulations using a 3D Cartesian box regular grid with 55 million computational points (275 million unknowns) and a 3D spherical shell fully irregular grid with 12 million computational points (60 million unknowns) and additional 240 million tracers. The strong scaling achieved with the two setups is shown in Fig. 1a. For the 3D Cartesian box grid we used up to 54 x 10^3 cores, while the 3D spherical shell simulation has been performed with up to 13 x 10^3 cores. The domain decomposition for the 3D spherical shell grid used 2 radial and 4352 lateral slices (4352 being the largest amount of available Thomson-points). Fig. 1b-d shows a typical domain decomposition for various meshes.

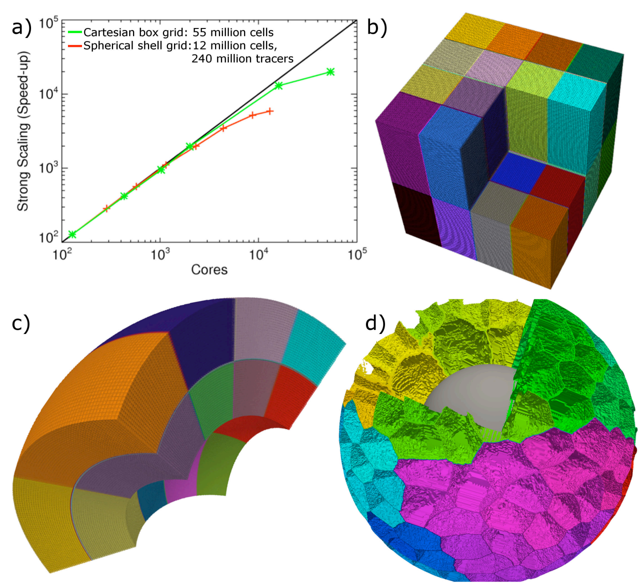


Fig. 1: a) Gaia code performance with up to 54 x 10^3 computational cores using a 3D Cartesian box grid with 55 million cells (green line) and a fully irregular spherical shell grid with 12 million cells (red line); Domain decomposition b) on 32 processors for a Cartesian box grid containing 1 million cells; c) on 16 processors for a partial sphere grid containing 1 million cells; d) on 255 processors for a fully irregular 3D spherical shell grid using a finer resolution towards the surface and a total of 8 million computational cells.

Application to mantle convection: *Magma ocean crystallization and onset of solid-state convection.* In the early stages of planetary evolution, the amount of heat available from accretion and core formation processes as well as from the decay of short-lived radioactive nuclides like ^{26}Al and ^{60}Fe is thought to have caused multiple episodes of extensive melting of the silicate mantle leading to local or even global magma oceans [9].

A better understanding of the solidification of a liquid magma ocean is essential for constraining the subsequent planetary evolution, but so far the crystallization process is poorly understood. For example, various studies predict a crystallization sequence of the Martian magma ocean that is difficult to reconcile with the subsequent thermochemical evolution and long-lived volcanic activity of the planet. Previous studies of magma ocean solidification have assumed an undisturbed crystallization process leading to a gravitationally unstable layering of the mantle that may cause an overturn [e.g., 10]. However, the density distribution attained during the crystallization process can be strongly influenced by the onset of solid-state convection prior to complete solidification. In a recent study, we investigated the mixing behavior in the Martian mantle during the crystallization of a global magma ocean and found that compositional heterogeneities established during the solidification can be partly or even entirely erased [11].

Thermochemical evolution of Mercury. We have applied our code to model the thermochemical evolution of Mercury and to predict crustal production and duration of magmatic activity that can be compared to the observations of the MESSENGER mission. This comparison led to inferences about the bulk abundance of radiogenic elements in the interior of the planet and the duration of volcanic activity [12]. The latter is compatible with the dating of the youngest large volcanic provinces [13]. Moreover, including the effects of large impact basins on the thermal evolution, we showed that it is possible to connect the local datasets relative to the large impact basins on the planet with its global thermal evolution, possibly providing a pathway to explain some of the geochemical anomalies associated with the interior of large impact basins [14].

Present-day surface heat flow of Mars. We have shown that the spatial variations of crustal thickness and the pressure dependence of the viscosity are responsible for the spatial distribution of surface heat flow and most likely affect the formation and location of mantle plumes [15]. Mantle plumes are directly related to partial melt production in the mantle and are thought to be responsible for the youngest volcanic activity in Tharsis and Elysium volcanic provinces. Nevertheless, our simulations predict that mantle plumes are unlikely to affect the upcoming surface heat flow measurements that will be performed by the InSight mission [16] and the heat flow values at the landing site in the Elysium Planitia region would be representative of the present-day average surface heat flow of Mars [15]. Moreover, using a number of constraints from the Martian geological record, (e.g., the evolution of the elastic lithosphere thickness, long-lived partial melt production in the mantle and geodetical estimates), along with a large set of numerical simulations, we can identify a best fit model of the Martian interior, which can be tested and validated with the upcoming seismic and heat flow measurements of InSight.

Conclusions and Outlook: Numerical simulations of interior dynamics can be applied to investigate the thermal evolution of a terrestrial body from the earliest stages up to the present day. Such models can be employed to interpret upcoming measurements in a global context and to link surface observations to the interior evolution. Moreover, with the increase of computational resources, the large amount of data produced by numerical thermal evolution models can be combined with machine learning algorithms and deep neural networks to identify key parameters that control the evolution of terrestrial bodies. This combination is a promising approach that has only started to be included within the scope of linking input parameters necessary to run thermal evolution models with available planetary mission data and observations.

References:

[1] Schubert G. et al. (2001) *Cambridge University Press.*[2] van Keken et al. (1997) *JGR*, 102(B10), 22,477–22,495. [3] Tackley P.J. and King S. (2003) *Geochem. Geophys. Geosyst.*, 4(4), 8302. [4] Plesa A.-C. et al. (2013), *IGI Global*, 302–323. [5] Hüttig C. and Stemmer K. (2008a) *Geochem. Geophys. Geosys*., 9(2), Q02,018. [6] Hüttig et al. (2013) *Phys. Earth Planet. Inter.*, 220, 11–18. [7] Hüttig C. and Stemmer K. (2008b) *Phys. Earth Planet. Inter.*, 171, 137–146. [8] Plesa A.-C. et al. (2015*) Springer International Publishing*, 675–687. [9] Solomatov S. (2001) Treatise on geophysics, 9, 91–119. [10] Elkins-Tanton L. et al. (2005) *Earth Planet. Sci. Lett.*, 236, 1–12. [11] Maurice M. et al. (2017) *JGR*, 122(3), 577–598. [12] Tosi N. et al. (2013) *JGR*, 118(12), 2474–2487. [13] Byrne P. et al. (2015) *GRL*, 43, 7408–7416. [14] Padovan S. et al. (2017) *Nat. Comm.* (in press). [15] Plesa A.-C. et al. (2016) *JGR*, 121, 2386–2403. [16] Banerdt B. et al. (2017) *LPS* XLVIII, abstract #1896.

SEABIRD: A FLEXIBLE AND INTUITIVE PLANETARY DATAMINING INFRASTRUCTURE. R. Politi[1], F. Capaccioni[1], M. Giardino[2], S. Fonte[1], M.T. Capria[1], D. Turrini[1], M.C. De Sanctis[1] and G. Piccioni[1], [1]INAF-IAPS Rome Italy (romolo.politi@iaps.inaf.it, Via Fosso del Cavaliere 100, I00133 Rome Italy), [2]ASI.

Introduction: Research activities on planetary science, from data analysis to theoretical modelling, need to deal with a large and extremely diversified amount of data. The heterogeneity of the data arises from the following factors:

- Data sources (Space missions, ground observations, laboratory experiments, numerical simulations);
- Data formats (PDS 3 and 4, Fits, ASCII tables, VO-compliant exchange protocols, etc.);
- Data types (spectra, images, iperspectral images, temperatures, etc.).

When searching for a specific target (e.g. specific planetary body or region) the difficulty of the task increases, because a deep knowledge is required to deal with different formats. The search for specific information in this sea of data is therefore becoming increasing complex and time expensive.

SeaBIRD: To simplify the data search process and to remove from the user the burden of handling the different formats and technical specifications of the data, we have developed the SeaBIRD (Searchable and Browsable Infrastructure for Repository of Data)[1] software and hardware infrastructure.

SeaBIRD Structure and implementation stategy: The SeaBIRD approach is the creation of a data space, called "atomic space", in which each single data unit is described by a metadata created by a semantic interpreter, starting from the original metadata, using a synonyms dictionary and a set of derivation rules. In this space, SeaBIRD clusters data by dimensions and defines a new "molecular space" where these dimensions, through suitable surjective mappings, are used to remap the clustered data. In the current release of SeaBIRD we translated this data description approach in a more complex and hierarchical data model generating a ORDBMS (Object Relational Database Management System) schema. Our implementation uses PostgreSQL with the GIS (Geographic Information System) extension.

SeaBIRD's Capabilities: SeaBIRD has been supplied with a web-based GUI (Graphical User Interface) to allow the users to easily search and access both the data and the associated metadata. SeaBIRD allows for two data mining approaches. On one hand, it allows the user to retrieve the original archived files containing the sought data ("*archive mining*" mode). In other hand, SeaBIRD can also provide the specific slice of data of interest for the user ("*information mining*" mode) repacked in a file with the desidered format (PDS, ASCII table/CSV, etc.).

The current version of SeaBIRD also allows to directly perform simple data manipulation tasks on the retrieved informations using a set of preimplemented primitive functions. These primitive functions can be combined by the user to create more complex workflows. The results of these manipulations can then be downloaded with the same approach used for the *information mining* mode.

All SeaBIRD's GUI features are developed using the newest Google™ visualization API to provide a powerful but inuitive interface.

SeaBIRD's API: We developed a SeaBIRD API (Application Programming Interface), whose interoperability software layer is based on the Django framework, to allow programs written in other languages to search and retrieve data from SeaBIRD using, either directly or indirectly, the HTTP (HyperText Transfer Protocol) protocol. The data are delived as JSON (JavaScript Object Notation) objects.

SeaBIRD's Data Readers: as a sub-product of its development process, SeaBIRD's data-reading modules are also being repacked into independent Python softwares to be released to the community. We currently released the VIRTIS-Venus Express reader, VIRTISpy[2].

SeaBIRD In Short: SeaBIRD main characteristics are:

1) easy and fast access to data;
2) personalized data slicing capabilities;
3) optimized data transfer volume;
4) solid and personalizable online data manipulation capabilities;
5) science-oriented interface that allows for abstracting data from archiving formats.

The current, fully-functional release of SeaBIRD is online and available to the community to explore the data provided by the intruments VIRTIS-Rosetta[3], VIRTIS-Venus Express[4] and VIR-Dawn[5], onboard the ESA missions Rosetta and Venus Express and the NASA mission Dawn respectively.

References:

[1] Politi R., Piccioni G. SeaBIRD - A VIRTIS-VEX Data repository, European Planetary Science Congress 2010,held 20-24 September in Rome, Italy, p.401. [2] https://github.com/VIRTIS-VEX/VIRTISpy [3] Coradini et al. VIRTIS: An Imaging Spectrometer for the Rosetta Mission (2007). Space Science Reviews, Volume 128, Issue 1-4, pp. 529-559. [4] Piccioni, G., et al., 2007. VIRTIS: The Visible and Infrared Thermal Imaging Spectrometer. ESA Special Publication, SP-1295, 1-27. [5] De Sanctis M.C., Coradini A., Ammannito E., Filacchione G., Capria M.T., Fonte S., Magni G., Barbis A., Bini A. and Dami M. The VIR Spectrometer, Space Science Reviews, Volume 163, Issue 1-4, p. 329-369

AUTONOMOUS SOIL ASSESSMENT SYSTEM: A DATA-DRIVEN APPROACH TO PLANETARY MOBILITY HAZARD DETECTION K. Raimalwala[1], M. Faragalli[1] and E. Reid[1], [1]Mission Control Space Services Inc. (1125 Colonel By Drive, 311 St. Patrick's Building, Ottawa, ON K1S 5B6; kaizad@missioncontrolspaceservices.com, michele@missioncontrolspaceservices.com, ewan@missioncontrolspaceservices.com).

Motivation: Planetary rover driving operations are conducted by strategic path planning using orbital assets while tactically accounting for hazards and science targets within the field of view of the rover. Current strategies include a variety of automated and manual processes to plan and adapt rover driving paths while leveraging a science and operations backroom to support tactical decision making. Tactically, science team members are engaged to provide daily advice to rover planners to determine safest routes using imagery downlinked such that new commands can be uplinked to the rover for the next sol [Arvidson et al. 2017]. The rover's actual drive path deviates from the planned route to avoid hazards and obstacles not evident in the map derived from orbital data. Pre-set thresholds for risks on distinct terrain classes, based on terrestrial testing and experience, are used to stop the rover's traverse and await further commands from the ground [1]. For example, slip thresholds are set where traverses are automatically stopped if slip exceeds a defined threshold. For Curiosity's traverse from Yellowknife Bay to Mount Sharp, a route was planned using HiRISE orbiter imagery to avoid bedrock-dominated slopes greater than 25°, sandy slopes steeper than 12.5°, aeolian dunes and ripples and rocks taller than one wheel diameter [2]. Figure 1 depicts the geomorphic map along Curiosity's traverse between Yellowknife Bay and the base of Mount Sharp.

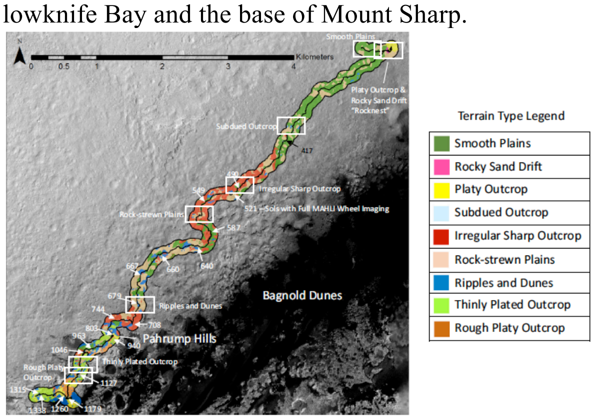


Figure 1: Geomorphic map of Curiosity traverse from Yellowknife Bay to Mount Sharp [1]

However, a limitation in this approach is that it is difficult to rapidly and accurately classify terrain types and hazards for tactical operations – particularly when encountering new terrain types. For example, Curiosity sustained significant wheel damage upon traversing rock-strewn terrain. As a result of punctures and tears, the remaining wheel lifetime is estimated at 10km motivating more careful path planning. To avoid further damage, Curiosity was directed towards megaripples (windblown, sand-sized deposit covered by coarser grains) to cushion wheel loads. However, the megaripples led to unexpected mobility difficulties, with high sinkage (approximately 30% of the wheel diameter) and high slip (up to 77%). Neither the rocky plateaus nor the megaripples were initially thought to pose mobility hazards to the rover by tactical operations teams and several sols were lost due to tactical planning and analysis.

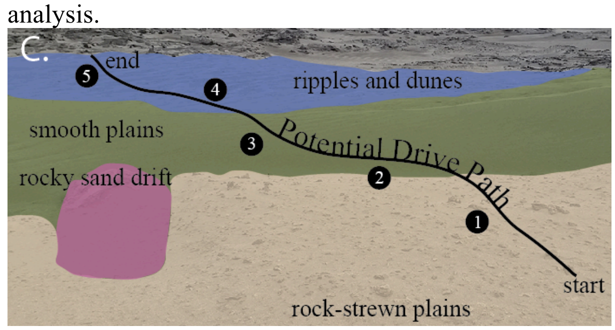


Figure 2: Image of classified terrain from Curiosity Navigation Camera (Navcam) on Sol 664 based on classes by [1].

The gap between rover navigation by human input and automated hazard detection and avoidance is filled increasingly by machine learning algorithms, which seek to make human-like geologic terrain observations [3,4]. These algorithms associate proprioceptive trafficability measurements and exteroceptive terrain sensing to make basic geologic terrain interpretations (e.g., sand vs. gravel) and predict terrain traversability ahead of the rover automatically. Because automated geologic terrain maps have not been extensively used in planetary rover operations, the efficiency and capability of operations workflows that incorporate such data have not been tested. We have tested this emerging protocol using the machine-learning terrain analysis software described below.

Autonomous Soil Assessment System Overview: The Autonomous Soil Assessment System (ASAS) is a software tool developed to predict non-geometric mobility hazards in the rover's field of view. ASAS uses a rover's navigation sensors to measure its mobility performance as it drives, and relates it to visual terrain features of the terrain. By fusing exteroceptive and proprioceptive data through machine learning algorithms, i.e. associating "feeling" with "seeing", ASAS learns to predict non-geometric hazards such as soft sand along the rover's path. In doing so, ASAS offers increased autonomy and efficiency in tactical workflows for rover navigation. Its three key components are:

Data-driven terramechanics modeling. Classical physics-based models have been used in the past to relate terrain mechanical properties such as cohesion and internal angle of friction to rover trafficability properties such as slip and sinkage. Such models rely on the accuracy of empirically derived soil parameters and are restrictive in structure, unable to capture unmodeled phenomena.

Our data-driven approach stores pairs of terrain slope and rover slip in finite queues and assigns heuristic hazard levels. This allows inference of trafficability from empirical data without strict model enforcement, offering a more robust solution over estimating soil properties and modeling wheel-soil mechanics. Additionally, this approach allows for online adaptability whereby the model is updated in real-time. Separate terramechanics models can be developed for discrete terrain classes; sand, gravel, and bedrock were chosen for our testing. Another advantage of the data-driven approach is that a new terrain class can be inferred from sufficient and consistent deviation of trafficability from existing models.

Automated terrain classification. Separating terramechanics models by terrain type requires the ability to classify that type. This is done by a supervised classification algorithm that uses images from the rover's stereo camera. Prior to running ASAS operationally, this algorithm is trained on images captured during a 'training phase', of terrain similar to what would be traversed in operation.

Real-time proprioceptive and exteroceptive data fusion and prediction. ASAS works by correlating what the rover sees to what it experiences, and then leveraging that correlation to predict what it would experience given what it sees ahead. Given a trained terrain classifier, ASAS identifies the terrain type within the field of view ahead of the rover. Once the rover reaches the area classified, with pose being estimated by localization techniques, ASAS stores the rover's slope angle (from its IMU) and instantaneous slip in the terramechanics model for that terrain class identified previously. This fusion of exteroceptive and proprioceptive information builds the data-driven terramechanics model over time. In operation, ASAS then identifies the terrain class and estimates the slope of an area ahead of the rover, and uses that information to predict rover slip and hazard level.

Operations, Field Tests and Demonstrations. ASAS has been developed to Technology Readiness Level 4 and field-tested at several sites in Ontario and Quebec, Canada, including the Canadian Space Agency's Mars Emulation Terrain, a.k.a. *Mars Yard*. Our platform is the Argo J5 built by Ontario, Drive & Gear and employs TRL 6 flexible metallic wheels. Additional field tests will take place at the dunes of White Sands, NM, in the winter of 2018, to test ASAS in a high-fidelity analogue environment consisting of aeolian landforms and duricrust-like features. Results from these tests will be presented at PSIDA-2018.

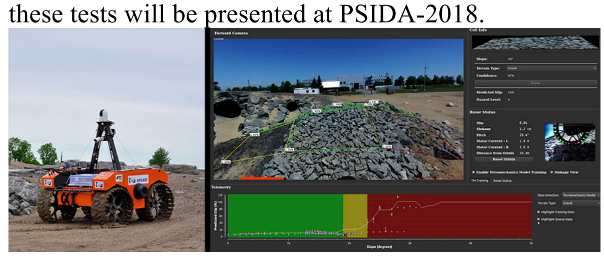


Figure 3: The Argo J5 rover and a screenshot of the ASAS user interface during a demonstration at the CSA Mars Yard.

Concept of Operations. Demonstrations of ASAS have involved a user interface through which an operator can monitor live information and train the terrain classifier and terramechanics model as the rover is driven manually in a low-latency teleoperation (LLT) environment (see Figure 3). The operator can use ASAS to inform their driving decisions based on hazards predicted ahead of the rover.

ASAS is designed to 'learn' to predict mobility hazards, thus improving the efficiency of tactical workflows in rover navigation and increasing the expected rover lifetime. It can be configured to be used as a path planning aid to suit an operation characterized by varying degrees of rover autonomy and command latencies. In the high-latency teleoperations (HLT) environment that characterizes Mars rover operations, ASAS incorporates non-geometric obstacle avoidance in autonomous navigation. The current strategy of stopping the rover when slip thresholds are reached can be augmented with ASAS' capability to predict that slip before the rover even reaches that hazardous terrain. For LLT of a lunar rover, an operator can use ASAS to aid their driving decisions in real-time.

References: [1] Arvidson R.E. et al. (2017) *J. Terramech., 73-93*. [2] Arvidson et. Al., (2014) *J. Geophysical Reaserch: Planets, 1322-1344*. [3] Gonzalez R. et al. (2017) *J. Field Robotics, 1-17*. [4] Wong C. et al. (2017) *IEEE Conference on Adaptive Hardware and Systems (AHS), 237-244*.

The PDS4 LDDTool – Information Modeling for Data Preparers. Anne C. Raugh[1] and John S. Hughes[2], [1]University of Maryland, College Park, MD, *araugh@umd.edu*, [2]Jet Propulsion Laboratory, Pasadena, CA, *John.S.Hughes@jpl.nasa.gov.*

Introduction: One of the major design goals of the PDS4 development effort was to provide an avenue for data preparers to extend the core PDS4 Information Model (IM) into their own contexts – specifically, to allow data preparers to define their own, new metadata values as needed to fully document their archive submissions. These local extensions, referred to as *data dictionaries*, must, of course, follow the same techniques, conventions, and restrictions as the core IM itself in order to be effective within the PDS4 system. Notwithstanding, expecting all data preparers to acquire expertise in information modeling, model-driven design, ontology, schema formulation, and PDS4 design conventions and philosophy is unrealistic, to say the least.

To bridge that expertise gap, the PDS Engineering Node has developed the data dictionary creation tool known as *LDDTool*. This tool incorporates the same software used to maintain and extend the core IM, packaged with an interface that enables a data preparer to create his contextual metadata using the same, standards-based framework as the PDS itself uses, and then output that metadata as the standardized set of schema files comprising a PDS4 data dictionary. Through this interface, the novice metadata developer has immediate access to the common set of PDS4 data types and unit classes for defining attributes, and a straight-forward method for constructing classes. The more experienced developer, using the same tool, has access to more sophisticated modeling methods like abstraction and extension, and can define very sophisticated validation rules. The generation of the PDS4-compliant schema files is handled entirely by *LDDTool*.

We present the key features of the PDS Local Data Dictionary tool (*LDDTool*), a tool that supports both the development of extensions to the PDS4 IM and ensures their compatibility with the IM.

The PDS4 Metadata Management System. Anne C. Raugh[1] and John S. Hughes[2], [1]University of Maryland, College Park, MD, *araugh@umd.edu*, [2]Jet Propulsion Laboratory, Pasadena, CA, *John.S.Hughes@jpl.nasa.gov.*

Introduction: The Planetary Data System (PDS) was established 30 years ago[1] as a permanent and living archive for the data returned by and relevant to NASA's planetary exploration program. The charge of the PDS was and is not merely to curate the data, but to maintain it in a usable state and make it available to contemporary scientists and researchers.

Metadata is the core element supporting all PDS activities. Structural metadata describes the physical format of the data, science metadata provides the analytical characteristics of the data, and provenance metadata established the history of the data. The essential problem of metadata is ensuring completeness and consistency among sources over time, while also allowing for contextual customization – so that data preparers can provide new metadata to document their specific circumstances.

The PDS system redesign known as **PDS4** has as its foundation the PDS4 Information Model – a codification of metadata for planetary data generally that also includes structures that can be used for the orderly and consistent extension of the IM into local data preparer contexts. We present the theory and methodology underlying the development of the PDS4 Information Model, the codification of metadata of different types into the PDS label structures, and the extension of that methodology into local contexts that can be defined by data preparers.

References:

[1] Lee, S. W. (1991) *The Planetary Data System, Revies of Geophysics Supplement, U.S. National Report to the International Union of Geodesy and Geophysics, 1987-1990*, 29, 337-340.

PDS4 CHALLENGES IN THE PSA. J. Saiz[1], I. Barbarisi[1], R. Docasal[1], C. Rios[1], A. Montero[1], A. Macfarlane[1], C. Laantee[1], S. Besse[1], C. Vallat[1], J. Marcos[1], J. Arenas[1], J. Osinde[1], C. Arviset[1], [1]ESA/ESAC, Camino Bajo del Castillo s/n, Urb. Villafranca del Castillo, 28691 Villanueva de la Cañada, Spain, jsaiz@sciops.esa.int.

Introduction: The Planetary Science Archive (PSA) [1] is the central repository where products from all planetary missions of the European Space Agency (ESA) are stored, following the standards given by the Planetary Data System (PDS).

While legacy missions such as Giotto, Huygens, Venus Express and SMART-1, the Rosetta mission, currently in post-operations phase, and the still operational Mars Express, use the former PDS3 standard, newer missions like ExoMars 2016, ExoMars RSP, BepiColombo and Juice use or will use PDS4.

Design challenges: Adopting PDS4 as the standard for new missions while being compatible with previously existing PDS3 products in the same archive has driven a design with several difficulties to overcome:

Common data model. The ESA planetary archive maps the key metadata from PDS3 and PDS4 into a common data model [2] with the intention of providing transparency to the available data lookup services: the main web portal from where products can be searched, viewed and downloaded, the machine access interfaces supporting the Planetary Data Access Protocol (PDAP) [3] and the EuroPlanet-Table Access Protocol (EPN-TAP) [4], as well as the FTP browser.

This common mapping is a result of a thorough analysis of the commonalities shared by both standards. Although fairly different in terms and format, their basic concepts are similar. An effort has been made to extract them into common categories (Figure 1).

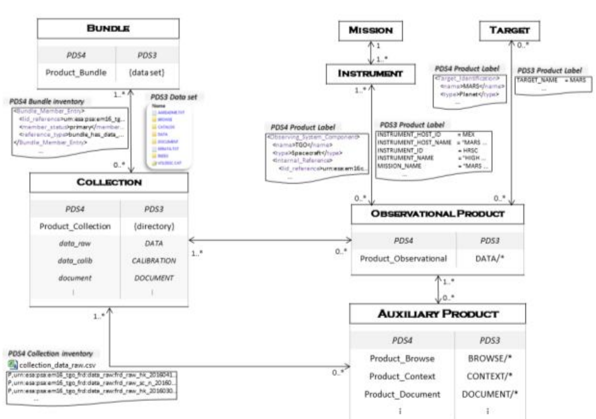


Figure 1. *PDS3 and PDS4 Data mapping*

Bundle generator. PDS4 products are organised at the top level into bundles [5]. Thus every delivery to the PSA coming from a PDS4 mission is expected to provide products within a bundle. Now, bundles have a non trivial structure to work with.

In order to make it easier for data providers to request products for being ingested into the PSA, it is permitted to deliver isolated products to the archive. Then it is the PSA ingestion software that is responsible for creating the required bundle for its proper ingestion (Figure 2).

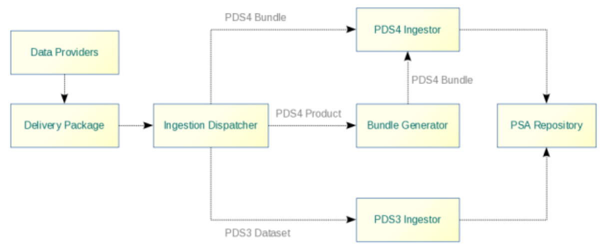


Figure 2. *Schematic view of the ingestion process in the PSA*

Creating a bundle and its delivery file, however, is not a straightforward operation, even for an automatic process, as the following paragraphs describe.

Validation procedure. Delivery files are expected to contain a transfer manifest, a checksum manifest and a label file, because in PDS4 every data item, including a delivery file, is considered a product, and each product has to be accompanied by a label file. All these have to be either validated, if a delivery file is passed to the PSA, or created for the generated bundle, if we talk about a single product delivery.

Subsequently, the data contents have to be validated too. This implies at least a syntax validation against the corresponding schema version, but also consistency checks to ensure that entities declared in the label files match the actual delivered contents, as well as verifying that values comply with *schematron* files.

The PSA will also delegate part of the validation process to software tools provided by the PDS that are yet to be incorporated.

Versioning of bundles, collections and products. Every time a given bundle is (re)delivered to the PSA, its version might be incremented. Bundle versioning is managed by the PSA so that it is kept consistent within the archive. The agreed convention with ESA missions employing PDS4 is to increase the minor version every week, and the major version every month. Therefore, when a bundle is going to be ingested, it is required for the ingestion process to look up and eventually retrieve its previous versions for a proper processing.

Deleting products considered erroneous or invalid by the data provider and reingesting them is a feature of the PSA that only applies to products that are not yet public. This task requires handling bundle and product versions with care in order not to create an undesirable maintenance problem.

Information Model (IM) versioning. Besides the aforementioned difference between PDS3 and PDS4 that the PSA has to manage, this latter standard has its own evolution. Roughly every six months, a new version of the PDS4 standard is released by the PDS. This implies a number of decisions to confront, both by the missions and by the archive:

- Which PDS4 version do we use?
- How often do we migrate to the last version?
- Do we upgrade all existing products?
- If each mission decides to use a different PDS4 version, how do we manage them in the PSA?

To facilitate this decision making to data providers, the PSA is building a flexible layout where various PDS4 versions are supported simultaneously: not only diverse information model versions between missions, but also different versions along the lifetime of a single mission. For this purpose, version independent Java interfaces have been created, which are implemented by adapter classes that make the appropriate translations to the corresponding schemas, with a minimal overhead. Though every time a new IM version is required to be incorporated, this common API could be slightly modified, the impact on client code is much less drastic than by exposing the generated JAXB [6] classes directly.

Ingestion time. As explained in previous paragraphs, creating a bundle involves many aspects, which may lead to a heavy and potentially slow processing. Consequently, it has been found that delivering fewer bundles with more products results in a more efficient strategy than trying to ingest numerous bundles with less products each.

Transactional ingestion. Once validated, adding a product or a bundle to the PSA comprises two clearly separated steps: ingestion in the data model (database), and importing the PDS4 files into the file repository. If either of the two fails, the whole operation has to be discarded, which has its own difficulty because, while transactions are nowadays well supported by database vendors, modifying and restoring the file system atomically is not so well provided by existing libraries.

PDS4 Updater. The received and imported PDS4 constructs contain all the information that providers have put in their delivery, but the data model reflects only part of them, a subset that is considered enough for the PSA interfaces to be exposed to users for searches and visualization purposes (the backed PDS4 data can be downloaded too, if necessary). This approach may lead to a situation in which the metadata stored in the database needs to be updated from the products of the repository, when for example a new key or some derived value is added to the data model and has to be filled from the existing imported files.

The alternative of reingesting these products would somewhat be an overreaction to this need. Therefore, it is foreseen to develop this functionality into the PSA to make it more tolerant to such update requests.

References: [1] Besse, S. et al., (2017) *Planetary and Space Science*; [2] Macfarlane, A. et al., (2017) *Planetary and Space Science*; [3] Salgado, J. et al. (2013), *IPDA Planetary Access Protocol*; [4] Erard, S. et al. (2014) *The EPN-TAP protocol for the Planetary Science Virtual Observatory*; [5] Data Design Working Group (2015) *PDS4 Concepts*; [6] E. Ort and B. Mehta (2003) *Java Architecture for XML Binding (JAXB)*.

DEVELOPMENTS IN GEOMETRIC METADATA AND TOOLS AT THE PDS RING-MOON SYSTEMS NODE. M. R. Showalter, L. Ballard, R. S. French, M. K. Gordon, and M. S. Tiscareno, SETI Institute (189 Bernardo Ave., Mountain View, CA 94043, mshowalter@seti.org, lballard@seti.org, mgordon@seti.org, rfrench@seti.org, mtiscareno@seti.org).

Introduction: The scientific analysis of most planetary data products requires a detailed, pixel-level understanding of the product's geometric content and timing. For spacecraft-based data, most scientists carry out such calculations using SPICE tools. At the PDS Ring-Moon Systems (RMS) Node, we have developed an object-oriented overlay on the SPICE toolkit, written in Python, that vastly simplifies and speeds up most of these calculations. With fully object-oriented underpinnings, "Object-Oriented Python/SPICE", or "OOPS", makes it possible to determine the geometric content of any supported data product with just a few lines of code.

Applications: OOPS provides the infrastructure to support many RMS Node activities, including:

- generating representative metadata to accurately describe the entire geometric content of data products in OPUS, our search engine;
- automating the navigation of Cassini images;
- conducting a variety of research projects involving images and other data products from Cassini, New Horizons, Voyager, Galileo, and HST.

Software Design: OOPS is designed around a large set of abstract object classes, each with a variety of implementations. For example, the abstract `Path` class defines an API that can be used to determine the position of one object relative to another as a function of time. One extension/implementation is `SpicePath`, which uses the SPICE toolkit to return this information. Alternative implementations describe linear motion (`LinearPath`), orbital motion (`KeplerPath`), etc. Similarly, the abstract `Frame` class describes a (possibly time-variable) coordinate system, with implementations including `SpiceFrame`, `RotatingFrame`, and `Cmatrix`. With this design, adding a new capability to OOPS simply entails coding up a Python subclass that implements the API required by the base class.

Other abstract classes in OOPS describe `Instruments`, `Observations`, fields of view (`FOV`), Observation timing (`Cadence`), etc. Adding support for a new instrument or data type in OOPS is just a matter of programming to the well-defined APIs for these classes. Once implemented, a single Python program could be used to analyze remote sensing data products regardless of whether they are 2-D images (framing, pushbroom, or raster-scanned), cubes, occultation profiles, or virtually anything else.

All OOPS operations handle arbitrary N-dimensional arrays using the NumPy extension to Python. A single call can, for example, calculate the line-of-sight geometry of every pixel in an image. Array operations occur quickly and, by all appearances, in parallel. The SPICE toolkit is normally too slow to handle ~ 10^6 calls, but OOPS handles this by making a reduced number of calls and then interpolating. This vastly speeds up the process while retaining full precision.

We are in the process of releasing OOPS as open source to the community. Interested individuals can contact the author for pre-release access.

Metadata: Most of NASA's planetary data sets are far too large to be used effectively without a mechanism for product-level search. Our product-level search engine, OPUS ("Outer Planets Unified Search", pds-rings.seti.org/search), allows for a very fine degree of granularity in searches based on diverse geometric constraints. For example, a search on images that captured the south pole of Enceladus at a high phase angle immediately returns the highest-quality plume images. OPUS's search capabilities are made possible because we pre-calculate the geometric content of each supported product using OOPS, and use that information to populate our databases. Tables of this same metadata can be downloaded from the node at pds-rings.seti.org/viewmaster/metadata for use in other applications.

Automated C-smithing: We are now completing a project to "C-smith" the C kernels for the entire Cassini tour. Our C-smithing yields a new C-kernel that is accurate to sub-pixel precision for most of the Cassini images. Our technique has been to generate model images using OOPS, which we then correlate with the actual data. The correlation function enables us to determine the pixel offset accurately. We have developed a variety of techniques for fitting to images of solid bodies (with and without visible limbs), atmosphere-enshrouded bodies such as Titan, rings (including non-circular features), and background stars with and without smear. With minor modifications, these same techniques could be extended to other planetary systems, missions, and data sets. If PDS could routinely provide C-kernels with pixel-level precision, we could eliminate the need for most scientists to perform image navigation as an (often tedious) first step in their data analyses.

MULTI-MISSION LASER ALTIMETER DATA PROCESSING AND CO-REGISTRATION OF IMAGE AND LASER DATA AT DLR. A. Stark, K.-D. Matz, and T. Roatsch, Institute for Planetary Research, German Aerospace Center (DLR, Berlin, Germany, Thomas.Roatsch@dlr.de)

Introduction: The Institute of Planetary Research (IPR) at DLR has a long tradition in the processing, analysis, and archiving of planetary mapping data. In the past years, we also have become interested in laser altimeter datasets from various planetary missions. Consequently, IPR has joined forces with Technical University of Berlin to develop a system for the processing and storage of laser altimeter datasets, which are typically very large. The system is based on a Linux server and uses PostgreSQL and PostGIS as database software. The system currently contains all data from the MOLA (Mars Global Surveyor), LALT (Lunar Kaguya mission), MLA (MESSENGER), and LOLA (Lunar Reconnaissance Orbiter) instruments. The main purpose of the system is a fast search for laser altimeter data points in a specified area on the surface of the planetary body. Additional functions for recovery of housekeeping data, as well as the calculation of physical parameters like slope and roughness are also implemented. It is planned to use this laser data processing system also for the planning and optimizing of the operations of upcoming instruments like BELA (BepiColombo) and GALA (JUICE).

Furthermore, laser altimeter data can be combined with digital terrain models based on stereo images (stereo DTMs). Through the co-registration the individual advantages of these complementary topographic datasets can be combined, while the disadvantages are avoided [1]. Typically, laser altimeter data provide high topographic accuracy but it suffers from large gaps in the coverage. For stereo DTMs typically the opposite applies, they provide extensive coverage but less accurate height information. Consequently the combination of both data sets leads to internally consistent topographic products and can be used for quality assessment. Moreover, the temporal coverage of the data sets can be used for measurement of the rotation and tidal deformation of the planetary object [2].

Data Ingestion: All laser altimeter data from previous missions (LALT, MOLA, and MLA) and from the currently running mission (LOLA) were downloaded from the Planetary Data System (PDS) nodes and ingested into the databases carefully looking for the quality of the different laser shots. The datasets were indexed and clustered using the PostgreSQL commands to allow very fast access to the data. Data from instruments on missions close to launch (BELA) and in preparation (GALA) were simulated using an instrument performance model by taking into account the individual instrument characteristics and operation scenarios [3].

Data Analysis: A set of PostgreSQL routines was developed to allow the analysis of the laser databases. Scripts for the calculation of e.g. topography, slope, and roughness were developed to allow the calculation using the fast algorithms which are inherent to PostgreSQL and PostGIS. One example of a roughness calculation from return pulse spreading is shown in Fig.1. The developed scripts can be used for various missions since the catalog structures were designed very similar. Recently, a Python interface to the database was established and allows a more convenient access to the databases.

Fig.1: Roughness plot of Candor Chasma calculated from MOLA data. White areas correspond to high roughness, while black areas indicate regions which are smooth. The baseline for the roughness measurements is 75 m, i.e. the diameter of the MOLA footprint.

Co-registration of laser and stereo image data: Due to limitation in knowledge of the spacecraft orbit and attitude the topographic datasets (laser profiles or stereo DTMs) can have offsets with respect to each other. With the help of the co-registration such offsets can be determined and both datasets can be brought to an agreement within their respective uncertainties [4].

Co-registration would be straightforward in the case one could easily identify conjugate points, i.e. measurements which correspond to the same feature on the surface. Due to the heterogeneity in the coverage of the laser profiles compared to stereo DTMs the identification of conjugate is very hard to establish. Thus, our approach consists of both finding conjugate points and determining the transformation of the co-registration.

Co-registration formalism: In particular our approach is to co-register points in 3-D to a quasi-continuous representation of the planetary surface. Thereby the former points are the laser altimeter measurements and the latter is given by the gridded stereo DTM with applied sub-pixel interpolation. Having a continuous representation of the surface has the advantage that the gradients for computation of partial derivatives can be easily obtained. Thus, the co-registration can be performed through a non-linear least-squares adjustment. The functional model g is formed by the radial differences of the DTM radius r^i_{DTM} and the i-th laser altimeter measurement $\boldsymbol{r}^i_{\mathrm{LA}}$

$$g^{\boldsymbol{i}}(\boldsymbol{p}) = r^i_{\mathrm{DTM}}(\boldsymbol{p}) - \left|\boldsymbol{r}^i_{\mathrm{LA}}(\boldsymbol{p})\right|, \quad (1)$$

where $\boldsymbol{p}$ is a vector of co-registration parameters. The stereo DTM $r_{\mathrm{DTM}} = r_{\mathrm{DTM}}(l, s)$ is represented by a structured grid of lines l and samples s obtained from spherical or Cartesian coordinates with the help of a map projection. The parameters of the co-registration can be any parameterizable type of a 3-D manipulation, e.g. a similarity transformation with parameters for translation, rotation and scaling. Furthermore, the co-registration parameters can be related to dynamical process, like rotation of the planetary object [1,2].

With the help of the surface gradients the partial derivatives $\partial g^{\boldsymbol{i}}(\boldsymbol{p})/\partial \boldsymbol{p}$ can be computed and combined in form of a design matrix $\boldsymbol{A}$. The best-fit parameters are then obtained through an iterative solution of the normal equation

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k - (\boldsymbol{A}_k^T \boldsymbol{A}_k)^{-1} \boldsymbol{A}_k^T \, \boldsymbol{g}(\boldsymbol{p}_k)\,, \quad (2)$$

where k denotes the iteration number. Typically after five iterations the estimates for the parameters converged and their formal uncertainty can be computed by

$$\boldsymbol{\Sigma}_p = \sigma_g^2 (\boldsymbol{A}^T \boldsymbol{A})^{-1}\,, \quad (3)$$

where $\boldsymbol{\Sigma}_p$ is the parameter covariance matrix and σ_g^2 is the variance of the final radial differences (Eq. 1). When applicable the observations can be weighted according to the accuracy of the laser altimeter measurement or the stereo DTM height uncertainty. With the help of the interpolation of the DTM the co-registration parameters can be determined accurate to the sub-pixel level for profiles with more than 100 data points on a rough terrain.

Fig. 2 shows a portion of the MOLA profile together with the stereo DTM based on images from the MarsExpress High Resolution Stereo Camera (HRSC) [5]. The co-registration revealed that to match the stereo DTM the MOLA profile required a shift by 23.8±5.9 m and 152.1±6.5 m in latitude and longitude directions, respectively. The radial offset is only 2.1±0.6 m.

Through the general construction of the code the co-registration method can be easily applied to laser altimeter data obtained from different instruments and DTMs with different resolution and coverage. The code is written in Python and includes a VICAR (Video Image Communication And Retrieval) Python interface for reading and saving DTM data. As outlined above the laser altimeter data is conveniently accessed through a connection to the PostgreSQL database. The co-registered data products can be linked with data from other instruments, e.g. spectrometers, and allow further investigations of the surface properties [6].

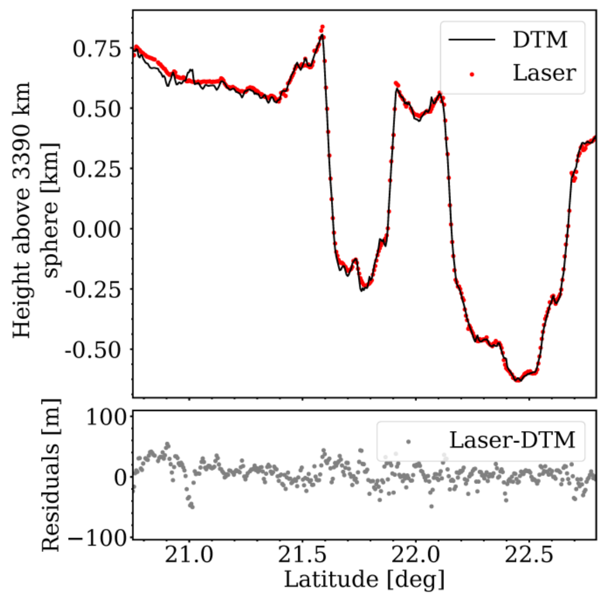


Fig. 2: MOLA profile (red dots) and HRSC stereo DTM (black line) located in the MC11 quadrangle [5] of Mars after co-registration. The residuals, which have a standard deviation of about 35 m, are shown in the bottom panel.

References:

[1] Stark, A., et al., 2015a. *PSS*, 117, 64-72. [2] Stark, A., et al. 2015b, *GRL*, 42, 7881-7889 [3] Steinbrügge, G., et al., 2015. *PSS* 117, 184-191. [4] Gläser, P., et al., 2013. *PSS*, 89, 111-117. [5] Gwinner, K., et al., *PSS*, 2016. 126, 93-138. [6] Naß, A. & D'Amore, M., 2018, this issue.

DEMONSTRATING THE OPEN DATA REPOSITORY'S DATA PUBLISHER: THE CHEMIN DATABASE. N. Stone[1], B. Lafuente[2], T. Bristow[2], A. Pires[3], R. M. Keller[2], R. T. Downs[3], D. Blake[2] , C. E. Dateo[2] and M. Fonda[2]. [1]Open Data Repository, Gray, ME (nate.stone@opendatarepository.org) [2]NASA Ames Research Center, Moffett Field, CA, [3]University of Arizona, Tucson, AZ.

Introduction: Recently, federal agencies have begun mandating that data and results from government funded scientific research be available and useful to the public and the science community. While large, homogenous fields often have repositories and existing data standards (e.g. GenBank), for small communities in multidisciplinary fields publishing and sharing data can be challenging.

In development for nearly four years, the Open Data Repository's (ODR) Data Publisher software has been designed as a collaborative data publication tool for small groups of independent researchers who usually have few options for publishing data that can be utilized within their community.

Objectives: ODR's Data Publisher aims to provide an easy-to-use software tool that will allow researchers to create and publish database templates and related data. The end product will facilitate both human readable interfaces (web-based with embedded images, files, and charts) and machine-readable interfaces utilizing semantic standards.

Characteristics: The Data Publisher software runs on the standard LAMP (Linux, MySQL, Apache, PHP) stack to provide the widest server base available. The software is based on Symfony (www.symfony.com) which provides a robust framework for creating extensible, object-oriented software in PHP. The software interface consists of a template designer where master database templates can be created and customized (Fig. 1). A master database template can be shared by many researchers to provide a common metadata standard that will set a baseline for all derivative databases. Individual researchers can then customize their instance of the master template with specialized fields, file storage, or visualizations that may be unique to their studies. This allows groups to create compatible databases for data discovery and sharing purposes while still providing the flexibility needed to meet the needs of scientists in rapidly evolving areas of research.

The platform facilitates flexible permission sets that enables researchers to share data collaboratively while improving data discovery and maintaining ownership rights. A web-based interface allows researchers to enter data, view data, and conduct analyses using any programming language supported by JupyterHub (http://www.jupyterhub.org). This toolset makes it possible for a researcher to store and manipulate their data in the cloud from any internet capable device. Data can be embargoed in the system until a date selected by the researcher. For instance, open publication can be set to a date that coincides with publication of data analysis in a third party journal.

A CSV import function will automatically generate a template and populate databases from a spreadsheet, allowing users to import large sets of data in a very short time.

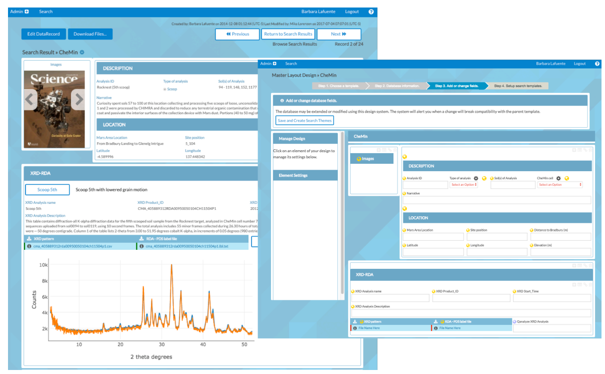


Figure 1. Example of data record view (left) and database template (right) from the CheMin Database.

The CheMin Database: In conjunction with teams at NASA Ames and the University of Arizona, a number of pilot studies are being conducted to assess the needs of individual research groups having disparate projects and data types and to guide the software development so that it allows them to publish and share their data collaboratively. These pilots include the CheMin Database (http://odr.io/CheMin), which contains the data products of the analysis performed by the Chemistry and Mineralogy (CheMin) X-ray diffraction instrument onboard the Mars Science Laboratory, together with tools and procedures for analysis of the data.

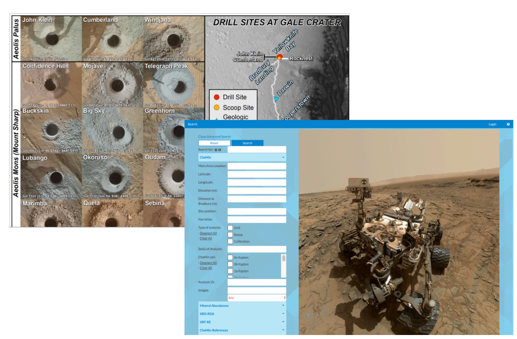


Figure 2. Search interfaces in the CheMin database.

The database benefits from the capabilities of the ODR software which provide a user-friendly interface, where the data are easily accessed using search tools (Fig. 2), visualization using a versatile graphing system, and data downloads in different formats.

The main goal is to provide outside users with the information and data analysis tools that are required to understand and re-analyze the original raw data, replicate experiments, or even perform entirely new studies with different starting hypotheses. Each data record includes: 1) sample description; 2) interactive XRD and XRF patterns with associated metadata and downloadable files; 3) mineral abundances derived from diffraction data; 4) access to the library of CIF files used in diffraction pattern analysis; 5) links to raw data and results from other MSL instruments (such as elemental composition data from APXS) for each of the samples analyzed by CheMin; 6) library of references associated to each analysis; 7) access to the Experiment Data Record (EDR) for each sample; 8) a detailed narrative of how the analysis was performed.

The database also provides access to *QAnalyze* (http://xrd.qanalyze.com/), an automated cloud-based application for quantitative analysis of mineral samples using X-ray diffraction (XRD) (Fig. 3).

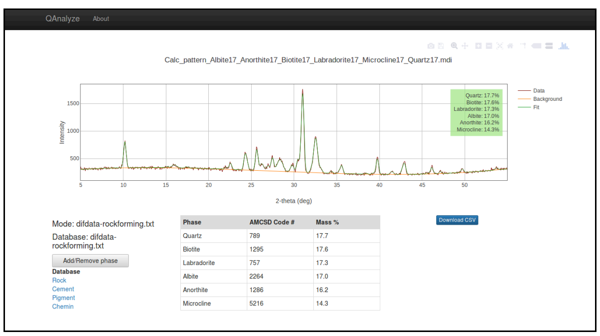


Figure 3. Example of analysis in QAnalyze.

Summary: A key feature of databases created using ODR is their ability to change and evolve over time. New data fields can be included and linked without disrupting the basic structure of the database, links can be created for new types of analyses and presentation formats. We are continually adding features and capabilities to the CheMin database (and other databases in the pilot ODR study) as they come available or are seen to be useful.

Acknowledgment: We gratefully acknowledge the support for this study by the Science-Enabling Research Activity (SERA) and NASA NNX11AP82A, Mars Science Laboratory Investigations.

ANALYSIS OF SIMULATED TEMPORAL ILLUMINATION AT THE LUNAR PSRS T. J. Thompson[1], P. Mahanti[1], The LROC Team [1]Lunar Reconnaissance Orbiter Camera, School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA (tthompson@ser.asu.edu);

Introduction: The Moon's lower range of sub-solar latitudes (-1.5° to 1.5°) through the seasons and longer days distinguish lunar lighting conditions from those of Earth. The absence of atmosphere means most of the incident photon flux impinging upon the Moon reaches the surface.

Some areas of the lunar surface never receive direct illumination (permanently shadowed regions, PSRs) [1] and have drawn illumination modeling interest [2], with emphasis on conditions for ice stability [3]. Furthermore, the absence of atmosphere also means the absence of light getting scattered by air, thus the common secondary illumination is from ground scattered light.

How bright are areas such as crater walls that can be seen from an arbitrary observer location, such as the bottom of a PSR? To answer this, we simulate lighting and only select areas visible to the observer. To further understanding of the topographic influence on the secondary illumination we calculate viewfactors [4] for all surfaces in the map. For our viewfactor maps, each cell value is the fraction of incident light at that cell which will go on to intersect another cell. This method requires DTMs, sub-solar points, and points of interest (receiver locations). So while this approach was designed with PSRs in mind, it may be used in other contexts.

Temporal Trends in Primary Illumination Within Line of Sight: We employ existing methods [5] to simulate surface lighting (cosine of incidence angle) and collect statistics for the set of pixels within the viewshed of an observer. First, sub-solar points are calculated using SPICE [6] through WebGeocalc (https://naif.jpl.nasa.gov/naif/webgeocalc.html) over a time period of interest. The illumination simulation program map_illum_layer [5] requires the sub-solar point to compute a solar vector. The angle between this vector and the surface normal is the incidence, i. The surface is modeled as brighter at smaller incidence angles, where the cell value, v is determined by $v = cos(i) \times 255$. Lunar Orbiter Laser Altimeter (LOLA) data [7] is used poleward of 79°latitude as the elevation basemap necessary for calculating shading. One map is created for each of the sub-solar points since the sun position changes with time. A viewshed map is generated for the chosen observer position and multiplied by the illumination map. Unlit pixels and those outside of line of sight are turned null. Summary statistics are calculated for the result of this multiplication. The data for all the output maps is stored as a table so these summary values may be compared across the time range (Figure 1).

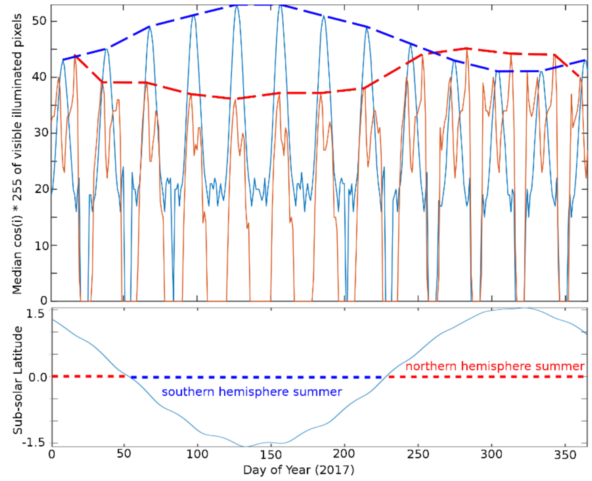


Figure 1: Seasonal relationship between visible, illuminated pixels from the perspective of a location on the floor of a north pole (red, Fibiger) and south pole (blue, Scott) PSR crater simulated for 2017

An example of summary output plots is found in Figure 1. The longest wavelength signal indicated by the dotted lines over the peaks (blue for Scott, red for Fibiger) shows the median value for the set of illuminated pixels within the line of sight of an observer on the floor. The largest median occurs during each pole's respective summer. The individual peaks and valleys trace the lunar day and night. Other peculiarities may be attributed to local topography. Figure 2 illustrates the illumination maps at different times in the lunar day for Fibiger crater, which hosts a north pole PSR.

Topographic Effects on Secondary Illumination: The orientation and position of a light emitter and receiver is key in the secondary flux at the receiver. A viewfactor is the fraction of the incident light on one object output towards another object. This method has been applied in a planetary context by Vasavada et al. [3] for lunar and mercurian craters for computation of thermal balance, but we are using the same principle in the context of secondary illumination in primary shadows. To calculate this coefficient (Eq.1, after [3]) we use the DTM to derive the positions and orientation of all potential emitters compared to the receiver. The DTM is first convolved with 3x3 filters to calculate x and y directional slopes for each pixel. The pixel scale and elevation is used to generate the x direction and y direction vectors which describe an oriented plane. Each

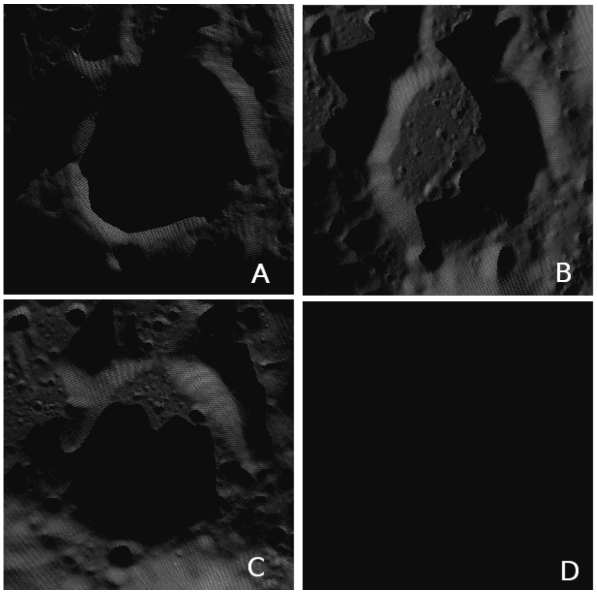


Figure 2: Illumination conditions at the northerly Fibiger crater at 7 day intervals in January 2017. A.) January 1st. B.) January 7th. C.) January 14th. D.) January 21st.

cell or plane in the raster is treated as an emitter. The geometry between the emitter and receiver surface normals and the line of sight between the two are calculated according to Eq 1. Planes that face each other and are close together have larger viewfactor coefficients.

$$a_{ij} = \frac{1}{\pi}\frac{(\cos(\theta_1)\cos(\theta_2)A_2)}{R_{ij}^2} \quad (1)$$

As the whole map contains a_{ij} values, it is then possible to multiply a map containing primary illumination values, I_{ij}, for each cell in the raster, and also an albedo map, q_{ij}, to account for reflectivity. After this element by element multiplication, summing all elements of the matrix gives the total flux at the specified receiver plate. This function calculated for all cells is outlined with equation 2.

$$I_{i_r j_r} = \sum_{i=0}^{n_r-1}\sum_{j=0}^{n_c-1} a_{ij} \times I_{ij} \times q_{ij} \quad (2)$$

The output a_{ij} map is stored as a 32 bit geotiff with the same offsets and pixel scales as the DTM. The a_{ij} values decrease with the square of distance from the receiver. Also, nearby terrain tends to have similar orientations inhibiting received scattered light, as surface normals need to be colinear and opposite for greatest light reception. The expected patterns for lunar craters are low amounts of incident scattered light in the flats due to near-parallel plate orientations and increased incident scattered light where the slopes of the inner wall of the crater are steep, then a fading with distance. Observing Figure 4, receiver 1 tilts towards the top of the

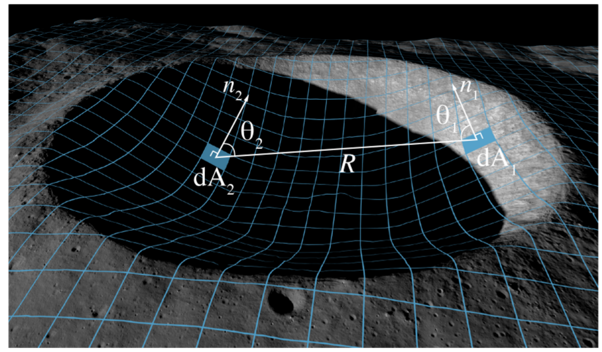


Figure 3: Angular relationship between emitting and receiving planes in a crater demonstrated with a digital model of Sylvester N crater

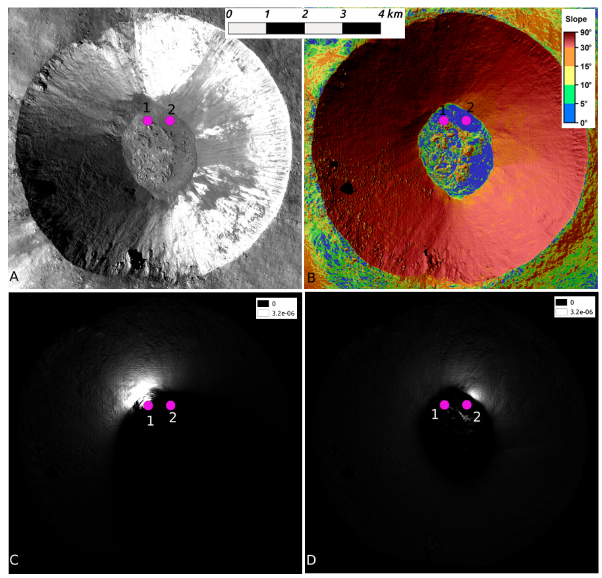


Figure 4: A. 90cm/px Eimmart A narrow angle camera (NAC) image, up is north. The pink dots are receiver 1 and receiver two. B. slope map. C. viewfactor map for receiver 1. D. viewfactor map for receiver 2. Source of imagery, topography, and slopemap from http://wms.lroc.asu.edu/lroc/view_rdr/NAC_DTM_EIMMARTA

image, so it is able to receive light from the floor in front of it, but since the receiver plate is tilted forwards, very little light reaches the receiver from the south. Receiver 2 is on flat ground, thus other flat areas contribute little light, it is the steepness of the nearby wall that contributes most scattered light according to the model.

References: [1] K. Watson, et al. (1961) *JGR* 66(9):3033. [2] E. Mazarico, et al. (2011) *Icarus* 211(2):1066. [3] A. R. Vasavada, et al. (1999) *Icarus* 141(2):179. [4] D. C. Hamilton, et al. (1952) . [5] C. Hanger, et al. (2013) in *Lunar Science Forum*. [6] C. Acton, et al. (2016) *ISPRS* 357–359. [7] D. E. Smith, et al. (2010) *Geophysical Research Letters* 37(18).

TOOLS TO MANAGE AND ACCESS THE NOMAD DATA. L. Trompet[1], A.C. Vandaele[1], I. R. Thomas[1] and the NOMAD team, [1] Royal Belgian Institute for Space Aeronomy (IASB), 3 av. Circulaire, 1180 Brussels, Belgium (loic.trompet@aeronomie.be).

Introduction: The NOMAD (Nadir and Occultation for MArs Discovery) [1] is one of the four instruments on-board the ExoMars Trace Gas Orbiter. It consists of three high-resolution spectrometers sensitive in a wide spectral range from IR to UV. It is also able to operate with different viewing geometries: solar occultation, limb and nadir. Thereby, the NOMAD instrument will extend the current database of species composing the Martian atmosphere.

The nominal mission of the ExoMars spacecraft will start in March 2018. Nevertheless some observations have already been performed to ensure that NOMAD works correctly. These early data are also used to start the data pipeline of the treatment processes that will generate different level of calibrated spectra. Complex algorithms will be used with the fully-calibrated spectra to derive temperature, pressure and density profiles of the atmospheric constituents of Mars. This last step requires significant computational means. Furthermore, the whole dataset generated from this pipeline will require important storage capabilities.

The three spectrometers of NOMAD will generate a huge amount of data to treat and share. The NOMAD team can rely on the knowledge acquired from the SOIR [2] instrument that was on-board the Venus Express spacecraft.

Treatment: The treatment of the data of NOMAD can be separated into two main steps: the calibration of the spectra and the profiles derivation. The first step requires little computing resources but the second one needs significant computing resources.

The ASIMUT radiative transfer code will be used for the profiles derivation. It allows to retrieve profiles following the different viewing geometries and the spectral range of NOMAD including consideration of absorption and scattering processes. The ASIMUT program is based on the Optimal Estimation Method developed by Rodgers (1990, 2000) [3,4]. An online version of this program is available at https://asimut.aeronomie.be/ and the ASIMUT documentation is available at http://planetary.aeronomie.be/en/asimut_documentation/html/index.html. ASIMUT can be used for spectra simulation and for profiles retrievals from a series of spectra obtained by different Instrument types under different viewing geometries. This is the reason why ASIMUT is a tool of particular interest for Virtual Observatories. Therefore some efforts are currently being made to make ASIMUT available through the Virtual Observatory VESPA [5].

The ASIMUT program is regularly improved. It will in particular be optimized to be applied on NOMAD observations. Algorithm to manage the different runs of ASIMUT will be used as it was used for processing data obtained by SOIR. Furthermore, when possible, some parts of the code will be parallelized and run on the Space Pole HPC from Brussels, Belgium.

Sharing: An overview database has been created containing information such as the date, relevant geometry, spectral range for each measurement. This database can be accessible through the NOMAD database interface at http://mars.aeronomie.be/en/exomars/obs-overview/.

The calibrated spectra will be available for downloading in the ESA PSA (as it is already the case for SOIR spectra) using the PSA psa.esa.int interface. The profiles, for their part, will be accessible through the VESPA research infrastructure as this is already the case for SOIR profiles.

Summary: The NOMAD instrument on-board the ExoMars Trace Gas Orbiter will generate a large amount of data of the Martian atmosphere. The Planetry Aeronomy Division at IASB is willing to make their tools (such as ASIMUT and the overview) and these data (using the PSA and VESPA) available to the whole planetary science community.

References:

[1] Vandaele A.C. et al. (2015) *Planet. Space Sci., 119,* 233-249. [2] Vandaele A.C. et al. (2016) *Adv. Space Res., 57,* 443-458. [3] Rodgers C.D. (1990) *J. Geophys. Res., 95(D5),* 5587-5595. [4] Rodgers C.D. (2000) *World Scientific, ed. N.J. Hackensack,* University of Oxford. [5] Erard S. et al. (2017) *Planet. Space Sci., In Press.*

MACHINE LEARNING APPROACH TO DECONVOLUTION OF THERMAL INFRARED (TIR) SPECTRUM OF MERCURY SUPPORTING MERTIS ONBOARD ESA/JAXA BEPICOLOMBO.
I. Varatharajan[1], M. D'Amore[1], A. Maturilli[1], J. Helbert[1], H. Hiesinger[2] [1]Institute for Planetary Research, German Aerospace Center DLR, Rutherfordstrasse 2, 12489 Berlin, Germany (indhu.varatharajan@dlr.de), [2] Wilhelms Universität Münster, Germany

Introduction: Spectroscopy is the powerful technique to study the surface mineralogy of any planetary body from its orbit. Spectrometers with wide spectral range, greater spectral and spatial resolution with repeated orbital coverage are helping us to map the surface mineralogy of planets in greater detail. Various spectral ranges tell different stories and properties of the surface we look at. For eg., VIS-IR spectroscopy for a rocky planet would tell us about the distribution of Fe,Ti,Mg,Ca rich minerals for both its igneous and sedimentary phases whereas thermal IR spectroscopy reveals the Si-O abundance on the bulk mineralogy of the pixel we look at. By carefully understanding the spectral behavior of various planetary analogues in laboratory experiments at the planetary surface and environmental conditions, one can map the mineral abundance and distribution globally from orbit.

Challenges in Spectroscopy: Many factors imparts changes to spectral behavior of a mineral such as grain size, phase angle of observation, slope, and abundance of the mineral. Though these factors affecting the spectra can be understood in a controlled environment, the real challenge comes in understanding the spectra can be addressed in two parts: 1. spectral behavior of minerals in their related planetary environment and 2. understanding the mixture spectra containing more than one mineral.

Planetary Spectroscopy Laboratory (PSL): Over the last 10 years the Planetary Spectroscopy Laboratory (PSL) located at the Institute of Planetary Research (PF) at the German Aerospace Center (DLR) in Berlin, Germany has been operating in various configurations to provide emissivity, reflectance, and transmission spectra of various rocks/minerals for the study of planetary and minor bodies surfaces [1,2,3,4].

PSL operates two identical FTIR (Fourier transform infrared) spectrometers (Bruker Vertex 80V); one spectrometer is equipped with aluminum mirrors optimized for spectral measurements in the ultraviolet (UV), visible and NIR (near infrared) wavelength region (say, 0.2 – 25 μm), and the second one is equipped with gold-coated mirrors optimized for measurements in near- to far-IR spectral range (1 - 100 μm) (Fig. 1). Both the spectrometers use a Bruker A513 variable-angle reflection accessory allowing biconical reflectance measurements under vacuum conditions for phase angles between 26° and 170° (Maturilli et al., 2014). The second spectrometer is also attached to an external chamber for direct emissivity measurements by heating the samples under vacuum to required temperature (~320 K – 1000 K) using a high efficiency induction heating system. The emissivity chamber is equipped with temperature sensors (thermopiles) for tracking the temperature of the sample and the surrounding environment during the measurements along with a webcam for monitoring the experiment. Thermal infrared spectral studies of a variety of mineral analogues to Mercury and other planetary bodies have been conducted in varying temperature conditions at PSL using this facility [5].

Figure 1. Laboratory set-up at PSL

Approach: The Mercury Radiometer and Thermal Imaging Spectrometer (MERTIS) payload of ESA/JAXA BepiColombo mission to Mercury will map the thermal emissivity at wavelength range of 7-14 μm and spatial resolution of 500 m/pixel [6]. Mercury was also imaged at the same wavelength range using the Boston University's Mid-Infrared Spectrometer and Imager (MIRSI) mounted on the NASA Infrared Telescope Facility (IRTF) on Mauna Kea, Hawaii with the minimum spatial coverage of 400-600 km/spectra which blends all rocks, minerals, and soil types [7] (Fig. 2). Therefore, the study [7] used quantitative deconvolution algorithm developed by [8] for spectral unmixing of this composite thermal emissivity spectrum from telescope to their respective areal fractions of endmember spectra; however, the thermal emissivity of endmembers used in [7] is the inverted reflectance measurements (Kirchhoff's law) of various samples measured at room temperature and pressure.

This compels us to re-examine the results by only considering the endemember spectra measured from simulated environment of Mercury.

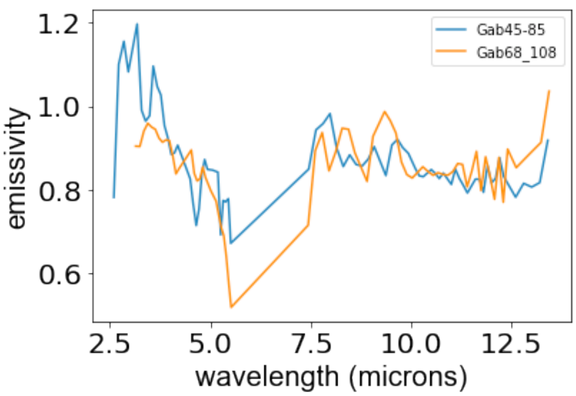


Figure 2. Telescopic spectra of Mercury [7]

Over a decade, the Planetary Spectroscopy Laboratory (PSL) facilitates the thermal emissivity measurements under controlled and simulated surface conditions of Mercury by taking emissivity measurements at varying temperatures from 100° to 500°C under vacuum conditions supporting MERTIS payload. The measured thermal emissivity endmember spectral library therefore includes major silicates such as bytownite, anorthoclase, synthetic glass, olivine, enstatite, nepheline basanite, rocks like komatiite, tektite, Johnson Space Center lunar simulant (1A), and synthetic powdered sulfides which includes MgS, FeS, CaS, CrS, TiS, NaS, and MnS. Using such specialized endmember spectral library created under Mercury's conditions significantly increases the accuracy of the deconvolution model results.

In this study, we revisited the available telescope spectra and redeveloped the algorithm by [8] by only choosing the endmember spectral library created at PSL for unbiased model accuracy with the RMS value of 0.03-0.04. Currently, the telescope spectra are investigated for its calibrations. Also, machine learning and Monte Carlo method is being studied for effective selection of endmembers from the large endmember spectral library of PSL and the results will be presented at PSIDA.

References: [1] Helbert and Maturilli, (2009) EPSL, 285 (3), 347-354. [2] Helbert et al. (2013a) EPSL, 369–370, 233–238. [3] Helbert et al. (2013b) EPSL, 371, 252–257. [4] Maturilli et al., (2008) PSS, 56 (3–4), 420–425. [5] Maturilli et al. (2017) LPSC, #1427. [6] Hiesinger, H. and J. Helbert (2010) PSS, 58(1-2): 144-165. [7] Sprague et al., (2009) PSS, 57, 364-383. [8] Ramsey and Christiansen (1998) JGR, 103, 577-596

Archive Inventory Management System (AIMS) – a fast, metrics gathering framework for validating and gaining insight from large file-based data archives. Rishi Verma[1], [1]Jet Propulsion Laboratory (4800 Oak Grove Drive, Pasadena CA 91108).

Abstract: The Archive Inventory Management System (AIMS) is a software package for understanding the distribution and characteristics of files and directories in large file-based data archives. It provides an efficient crawling and customizable extraction system for scanning over very large file directory trees and extracting specific information from assets within those directory trees. This information is then indexed in a search engine to provide robust analysis and visualization support in order to better understand characteristics of the archive.

The motivation for AIMS stems from a need within NASA's Planteary Data System (PDS) Imaging Node (IMG) [1] to *continually* keep up-to-date about file-based data archive characteristics, such as: adherence to file path naming conventions, file and directory sizes, checksums, validation of appropriate file placement within directories, etc. Given PDS IMG stores at least 63 TB of data and upwards of approximately 650 million file assets, keeping up-to-date with respect to these assets on a continual basis is a significant computational challenge. Existing approaches of sequentially processing each file asset can take weeks to finish, given the current PDS IMG archive size. To meet this computational challenge, as well as future ones, the AIMS system is being designed to take advantage of cluster computing, using the Apache Spark framework [2], and clusterized search tools to aggregate metrics metadata using the Elastic Search framework [3]. Taking a cluster-computing approach has the advantage of horizontally scaling and distributing the processing work of evaluating each file data system asset across a potentially increasing number of worker nodes. Thus, hardware purchases, not software design, is the limiting factor in reducing overall archive scanning and metrics extraction time.

Extracting and collecting metrics data is the first computational challenge AIMS seeks to address. Following this, AIMS also supports ad-hoc queries of collected metrics data for analysis purposes. For example, AIMS typically collects the directory size information of every directory within the PDS IMG archive; however, it also supports actions like summing up the total size of all the directories matching a particular directory path. To do this, the Elastic Search framework is utilized to provide ad-hoc, text-based and numeric aggregation support. In other words, an unforeseen query to the AIMS search engine results in a distributed computational search job where existing metrics are aggregated using the assistance of multiple cluster nodes to produce the final result. Thus, AIMS is flexible enough to provide very robust and efficient analytics capability that data scientists need to understand the characteristics of a file based archive. Together, with an efficient, distributed metrics extraction framework and a distributed, ad-hoc query capability, the AIMS system provides a fast and effective way to keep up-to-date with changing characteristics of file-based data archives.

References:

[1] "Cartography and Imaging Sciences Node" National Aeronautics and Space Adminstration. Last accessed: Nov 14th, 2017. <https://pds-imaging.jpl.nasa.gov/>.

[2] "Apache Spark" Apache Software Foundation. Last accessed: Nov 14th, 2017. <http://spark.apache.org//>.

[3] "Elasticsearch" Elasticsearch. Last accessed: Nov 14th, 2017. <https://www.elastic.co/products/elasticsearch/>.

PITSCAN: COMPUTER-ASSISTED FEATURE DETECTION. R. V. Wagner and M. S. Robinson. School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-3603.

Introduction: The Lunar Reconnaissance Orbiter Camera (LROC) consists of a single Wide Angle Camera (WAC) and twin Narrow Angle Cameras (NACs) that provide multispectral and high-resolution imaging, respectively [**1**]. The NACs are capable of acquiring panchromatic images at 0.5 m/pixel from an altitude of 50 km. A typical NAC image consists of 5,064 samples and 52,224 lines, resulting in over 500 megapixels of image data in each observation pair. As of 1 November 2017, the NACs have collected over 700,000 image pairs of illuminated terrain (~350 *tera*-pixels of data) covering most of the Moon. This dataset is far too large to search by hand, so we developed a feature detection tool (*PitScan*) to enable the discovery of lunar pits by pre-processing images to extract potential pits for human analysis[**2**].

Finding Lunar Pits with *PitScan*: Lunar pits are deep, vertical-walled collapse features, generally <100 m in diameter, and usually have an inward-sloping rim (**Figures 1**, **2**). So far, we have located over 300 pits across the lunar surface. Thirteen of the known pits are in mare flood basalts, three are in highland terrain, and the remainder are in young crater impact melt ponds. Their small size and rarity make pits a prime candidate for automated searching, and in fact most of the known pits were found using *PitScan* (previously reported in [**2,4**]).

Theory. Since the majority of slopes on the Moon are below the angle of repose (~36°) [**3**], very few features cast shadows when the Sun is within ~54° of the zenith. Pits, boulders, and other features with vertical surfaces all cast shadows at even lower angles; thus, a catalog of shadows in a high-Sun image should contain any pits in the area. All that is needed is to filter out pit shadows from non-pit shadows.

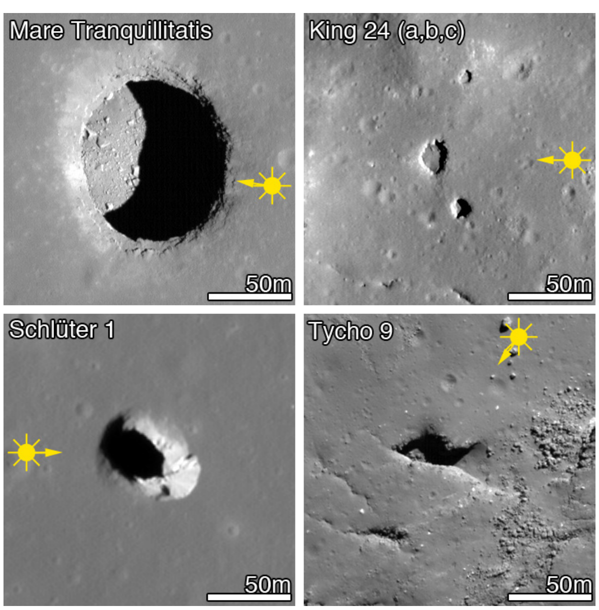


Figure 1: Examples of pits, demonstrating the range of sizes and morphologies that must be detected.

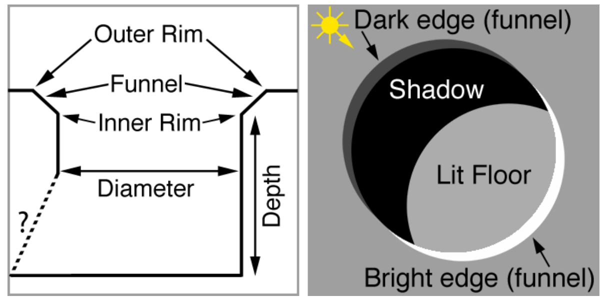


Figure 2: *Left:* Cross-sectional sketch of an idealized pit. Dashed line and question mark indicate possibility of overhang or cave entrance. *Right:* Overhead view of an idealized pit, showing how pit morphology produces distinctive features (compare to pits in **Figure 1**).

Implementation. PitScan was developed to locate all shadows larger than 15 pixels across (approximately the smallest size at which a pit can be visually confirmed), and exclude those features that are most likely to be boulders. The remaining potential pits are saved as small image clippings for a human analyst to check manually. *PitScan* runs on 16-bit calibrated image files (CDRs), and can complete a search of a single 250 megapixel NAC image in thirty seconds.

To find shadows, *PitScan* uses an empirically derived equation to calculate a cutoff value for "shadowed" pixels. The formula for this cutoff value (given image mean value μ, cutoff value $T = \mu \times 0.113 + 20$ [**2**]) was determined by manual inspection of pits in several dozen calibrated images with various Sun elevations. A complicating factor in tuning this cutoff is that pit interiors often have strong secondary lighting from sunlight reflected off of an illuminated wall, and thus can be brighter than shadows cast by rocks.

Once all the continuous blobs of pixels with I/F values below the cutoff that are at least 15 pixels across have been located, *PitScan* extracts a profile across each blob parallel to the solar azimuth, extending approximately 30 pixels beyond the bounds of the shadow (**Figure 3**). If the average I/F value on the up-Sun side of the shadow is greater than 0.9× the average value on the down-Sun side (a factor chosen to include most known pits in the images available when the algorithm was written), then the feature is assumed to be a rock, and discarded. For the remaining cases, *PitScan* saves a 300 × 300 pixel clipping for human review, along with a plot of the pixel values in the profile (similar to the left and right columns in **Figure 3**).

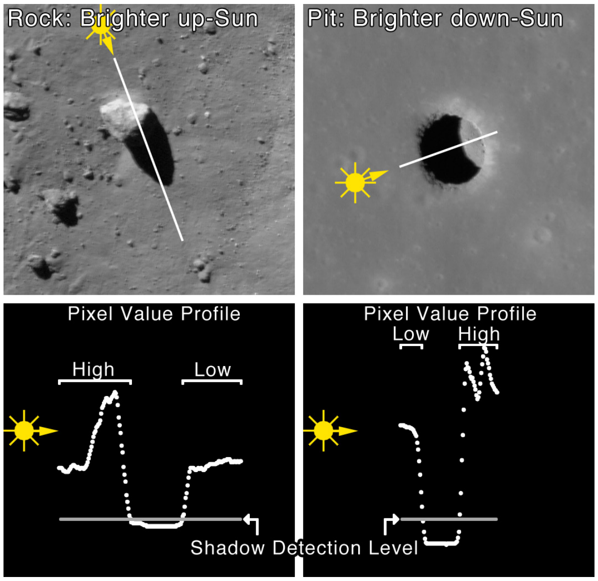


Figure 3: *Top*: Profile lines across detected shadows for rock (*left*) and pit (*right*). *Bottom*: Pixel values along the profile.

In cases where more than 50 potential pits are found in an image, *PitScan* instead saves a preview of the entire image with potential pit locations marked, so a human can check these feature-rich images for known patterns of false positives (such as outcrops on crater walls).

PitScan is normally used only on images with incidence angles (angle between Sun and zenith) less than 50°, as above this value shaded crater walls are frequently flagged as shadows, producing an excessive number of false positives.

Results. Excluding the feature-rich images, the algorithm historically has generated ~150 false positives for each successful pit identification. We consider this an acceptable level of false positives, as an experienced analyst can evaluate most image clippings in less than a second, and it only takes a few hours to check the results for six months of NAC images. The false negative rate can be greatly reduced by adjusting the ratio used to detect boulders: If the down-Sun side is at least 1.1x the up-Sun side, 40% of false negatives are excluded, and 85% of true positives are retained. To take advantage of this effect, the output clippings filenames start with this ratio, so the analyst can decide whether to start with the hundred or so clippings most likely to contain pits, or run through the thousands of false positives (and maybe one or two pits) at the other end of the list.

In a sample of all images from 2009-2016 of known pits with pixel scales such that the pit is at least 30 pixels across, and incidence angles < 50°, *PitScan* only detected 45% of the expected pits. Detection was better for non-impact-melt pits, with 86% of expected detections made, although there may be a sampling bias here, as 13 of the 16 known non-impact-melt pits were originally found using *PitScan*, while many impact melt pits were found by manual search near impact melt pits and related features identified by *PitScan*. Future work will focus on determining why pits are missed by the algorithm, and altering it to increase the detection rate.

Due to the limit on valid Sun elevations, *PitScan* can only search the region within ~50° latitude of the equator (77% of the Moon). To date, the NAC has acquired 307,824 images within the 50° incidence angle constraint, covering 76% of the searchable area.

Conclusion: *PitScan* allows for rapid searching of very large amounts of data to find small unusual features (pits). Currently we run *PitScan* 2-4 times per year (once or twice during each high-Sun imaging period) on all <50° incidence images acquired since the last run. The most recent run searched 17,907 images, using ~1,900 CPU-hours (not adjusted for CPU utilization- most of this time was likely spent on data transfer) on a 600-core processing cluster, and the output (~3,600 clippings) took about 1-2 hours of human analyst time to sort into pits, non-pits, and other interesting items (such as impact melt fractures and flow features, anomalously dark rocks, and a certain class of small, rocky craters with melt ponds). This run found eight images of pits, most of which were already known.

Pit discovery rates have fallen since 2012, mostly due to fewer new discoveries of pits in impact melt (mare pit discoveries have been relative constant, averaging ~1 per year). This drop-off is likely due to one of two factors: 1) The higher-altitude orbit LRO entered in late 2011 (leading to lower resolutions north of ~40° S, thus often placing the relatively smaller impact melt pits below *PitScan*'s size cutoff), or 2) most large young craters (the population most likely to have impact melt pits) having already been imaged early in the mission due to their scientific value.

Additional observations acquired during the LRO Cornerstone Mission and future extended missions will enable the LROC team to increase the spatial coverage of NAC images, and we will continue to use *PitScan* to find pits in this newly-imaged territory.

References: [**1**] Robinson, M.S. (2010), *Space Sci. Rev.* 129, 391-419. [**2**] Wagner, R. V. and Robinson, M. S. (2014), *Icarus*, 237C, 52–60. [**3**] Wagner, R.V. et al. (2013) 44th LPSC, #2924. [**4**] Wagner, R.V. et al. (2017) 3rd Planetary Data Workshop, #7074

USING AGISOFT PHOTOSCAN TO COMPARE TERRESTRIAL AND PLANETARY VOLCANIC FEATURES. R. V. Wagner[1], M. R. Henriksen[1], M. R. Manheim[1], M. S. Robinson[1], [1]School of Earth and Space Exploration, Arizona State University, 1100 S. Cady, Tempe, AZ 85287 – (rwagner@ser.asu.edu).

Introduction: The Agisoft Photoscan software aligns images from multiple cameras and uses shape-from-motion (parallax) to create high-resolution digital terrain models (DTMs) [**1**]. It is commercially available, and can ingest photos from common consumer- or professional-grade handheld cameras, in addition to more mobile imaging systems (i.e. drones). We have tried some similar open-source programs, but none of them allowed as much manual control. We created several high-resolution DTMs of terrestrial volcanic features that are believed to be analogous to features on the Moon and on Mars. We explore the potential for using these DTMs to better understand these analogs by comparing profiles and roughness measurements.

Terrestrial DTMs: We created DTMs of three terrestrial cones (summarized in **Table 1**) in Photoscan using photos taken with some combination of smartphones, handheld semi-professional cameras, and a DJI Mavic Pro drone (all of which used 12-16 megapixel sensors). The two Arizona cones were chosen for geographic convenience, while the Hawaii cone was selected as an Earth analog for lunar volcanic features. The photos for each site were taken by 1-3 operators. To minimize changes in lighting, images at two of the three sites were taken over 1-2 hours; at Colton crater, images were taken over the course of a 4.5 hour hike around the crater (although this did not have an obvious effect on the resulting model). The photos were imported and aligned with spatial control information derived from the geolocation tags embedded in all images except those from one of the handheld cameras; from the aligned images, we created a point cloud, which was then used to generate the final DTM. Processing each DTM took ~2-4 hours of human interaction and ~1-3 days of computer time.

Our previous analysis indicated that scale errors of ~2-3% are expected in the absence of high-accuracy ground control points [**2**]. A physical scale bar in the SP model had a length error of ≲0.5%. In the Colton model, ground points from a consumer GPS (3-4 m reported uncertainty) suggested a ~2° clockwise rotation and a ~1-2% scale error. We have not yet done error analysis on the Hawaii model.

Planetary DTMs: We compared our terrestrial DTMs with DTMs of features on Mars and on the Moon, created, respectively, by the HiRISE team [**3**] and the LROC team [**4**], with pixel scales from 1 m to 5 m (**Table 1**). In addition to volcanic cones, two Copernican impact craters of a similar diameter to Colton Crater (terrestrial maar) were also included to see if they were quantifiably distinguishable from the volcanic features.

Analysis: We compared profiles and depth/diameter ratios of these features to show that Photoscan DTMs can be used as analogs to planetary DTMs.

Background. Previous studies have used DTMs, both planetary and terrestrial, to investigate the ages and internal structures of volcanic cones. Small lunar cones (≲2km) have been considered similar to terrestrial cinder cones in terms of morphology; thus, our understanding of terrestrial formation processes has been used to interpret planetary cones [**5,6,7**]. Many authors [e.g. **7,8,9**] use DTMs to discuss differences in morphology, morphometry, and probable composition between cones on Earth, the Moon, and Mars, as well as implications for various models of small cone formation. Higher resolution DTMs on all bodies allow us to investigate the structure and composition of these cones and infer the eruption conditions that formed them. We now have high resolution DTMs on the Moon and on Mars [**3,4**]; the Agisoft Photoscan software allows us to quickly and conveniently create terrestrial DTMs of analogous features in order to make comparisons.

Methodology. Porter 1972 [**10**] characterized over 300 0.1-1.3km diameter cinder cones on Mauna Kea to establish typical ratios between cone width and depth, and between crater dimensions and cone size. Crater depth/crater width was found to be 0.14; however, this value represents a minimum depth only due to post-eruption filling, which can make them seem similar to terrestrial impact craters. Because our Photoscan DTMs are currently most complete over craters rather than entire cones, we used profiles taken over our 15 features (**Fig. 1**) to find crater depths and widths.

Comparison. We plotted crater depth and width for two terrestrial cones and one terrestrial maar, five martian cones, five lunar cones, and two lunar impact craters. **Fig. 2** shows that the terrestrial features were evenly distributed around [**10**]'s 0.14 ratio, while the lunar and martian cones had lower ratios, representing shallower/broader morphologies. The two impact craters fell above the line.

Conclusion: Photoscan provides a simple method of quickly producing very high-resolution DTMs of features, without requiring expensive and logistically difficult aerial photographic or LIDAR surveys.

The use of a drone in creating the model of SP crater significantly increased the quality of the DTM, by providing coverage of the flanks and a stable baseline of GPS coordinates, but was not necessary for modelling the interior of the crater. We found that modelling generally works best when there are images taken from within ~60° of the surface normal, so depressions can easily be mapped using handheld cameras. Flat surfaces and large positive relief features greatly benefit from the use of a low-cost drone to provide context and potentially tie together higher-resolution ground-level photography.

In addition to Earth analog sites, Photoscan can be used to model extraterrestrial surfaces using images from landed missions. Examples include route-mapping using images from the Mars Curiosity rover [**11**], and modelling the sampling trench dug at Shorty crater during the Apollo 17 mission [**2**].

References: [**1**] www.agisoft.com [**2**] Wagner et al. (2017) *3rd Planet. Data Wkshp.* #7023 [**3**] Kirk R. L. et al. (2008) *J. Geophys. Res., 113,* E00A24. [**4**] Henriksen M. R. et al. (2017) *Icarus 283,* 122-137. [**5**] Wood (1979) *Proc. Lunar Sci. Conf. 10th*, 3, 2815-2840 [**6**] Whitford-Stark and Head (1977) *Proc. Lunar Sci. Conf. 8th*, 3, 2705-2724 [**7**] Lawrence et al. (2013) *JGR Planets*, 118, 615-634, doi:10.1002/jgre.20060 [**8**] Stopar (2016) PhD dissertation, ASU [**9**] Brož et al. (2015) *JGR Planets*, 120, 1512-1527, doi:10.1002/2015JE004873 [**10**] Porter (1972) *GSA Bull.*, 83, 3607-3612 [**11**] Ostwald and Hurtado. (2017). *48th LPSC* #1787

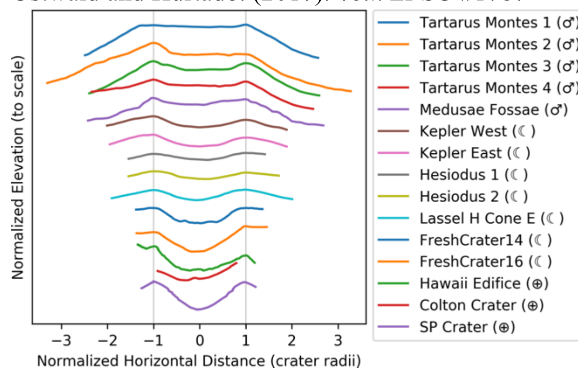


Figure 1: DTM profiles, normalized by peak crater radius. See **Table 1** for actual depths and diameters.

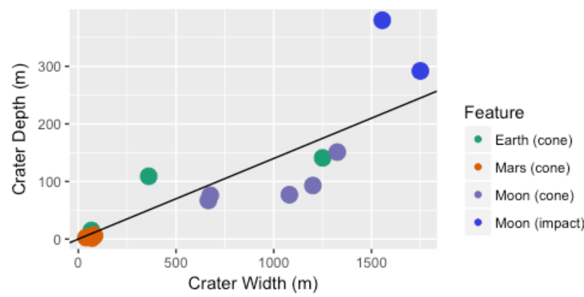


Figure 2: Plot of crater depth and width for the measured cones and impact craters. Line marks the 0.14 depth/diameter relationship found in [**10**] for terrestrial cinder cones.

Table 1: Terrestrial and Planetary DTM sites

Site	Planet	Camera(s)	Image Count	Scale (m/px)	Crater Widths (m)	Crater Depths (m)
Colton Crater	Earth	iPhone 7	52	0.23	1250	93
SP Crater	Earth	DJI Drone, iPhone 7, LUMIX DMC-LX100	434	0.12	360	109
Hawaii Edifice	Earth	Samsung Galaxy, Canon EOS 70D	496	0.08	65	15
Cone E of Lassell H	Moon	LROC NAC	-	3	1325	151
Fresh Crater 14	Moon	LROC NAC	-	3	1750	292
Fresh Crater 16	Moon	LROC NAC	-	5	1555	380
Hesiodus (two cones)	Moon	LROC NAC	-	3	1080, 1200	77, 93
Kepler (two cones)	Moon	LROC NAC	-	5	665, 675	67, 76
Cone near Medusae Fossae	Mars	HiRISE	-	1.01	40	2.6
Cones in Tartarus Montes (four cones)	Mars	HiRISE	-	1	70, 70, 80, 80	1.7, 5.3, 6.8, 5.3

DATA ARCHIVES AND STANDARDS FOR JAPANESE PLANETARY MISSIONS. Y. Yamamoto[1,2], Y. Ishihara[1], and S. Murakami[1], [1]Japan Aerospace Exploration Agency (3-1-1 Yoshinodai, Chuo-ku, Sagamihara, Kanagawa 252-5210, JAPAN). [2]SOKENDAI, The Graduate University for Advanced Studies, School of Physical Science, Space and Aeronautical Science (3-1-1 Yoshinodai, Chuo-ku, Sagamihara, Kanagawa 252-5210, JAPAN).

Introduction: Considering the long-term preservation of data, it is important to follow standards. Although the history of planetary explorations in Japan is over 30 years, standardization of data archives has just begun(Table 1). In the 1990s, Japanese planetary exploration programs did not notice the importance of data archiving. Recently discussions of data archives become active, and long-term preservation is one of the topics in the planetary exploration programs. There are several standards to be followed making planetary data archives.

Archiving standards: Japan Aerospace Exploration Agency (JAXA) has applied the Planetary Data System (PDS)[1] as a standard to the data archives of planetary explorations. The PDS is not just a format definition, it is the system as its name. It specifies directory structure and essential files to be provided.

The missions launched before 2011 uses PDS version 3 (PDS3), and after the Hayabusa2 mission, JAXA is going to use newly developed PDS version 4 (PDS4). The PDS4 incorporates recent Information Technologies, and it is superior to PDS3 regarding metadata verification and data updating mechanism. However, the conversion from PDS3 to PDS4 is under consideration because it requires too many resources.

Ancillary data standards: SPICE[2] is used to archive ancillary data such as trajectory, attitude, coordinates, shape model, etc. The first use of SPICE in Japan was the original Hayabusa mission. Through the experience of the past planetary missions and strong support of the JPL NAIF team, JAXA continuously accumulates knowledge on SPICE.

JAXA developed software to convert from the original format to SPICE format not to depend on a specific person. The software is included in an automated pipeline. It converts from a time conversion table to SPICE SCLK format, and from quaternions in the raw CCSDS space packets to SPICE CK format.

Coordinate and Cartographic standards: It requires prior adjustment of the coordinate standards. Also, projection rules should be determined in advance if possible.

An example is the SELENE(Kaguya) mission. It uses the mean Earth/polar axis (ME) reference system, and a reference sphere of a radius is 1737.4 km. There was a discussion to select a coordinate, and it decided to match with the Lunar Reconnaissance Orbiter (LRO) standards[3].

Table 1 Data of Japanese planetary explorers

Name	Launch	Target	Archives
Sakigake Suisei	Jan.1985	Comet Halley	DARTS PDS/SBN
Hagoromo Hiten	Jan. 1990	Moon	-
Nozomi	Jul. 1998	Mars	-
Hayabusa	May 2003	Asteroid Itokawa	DARTS PDS/SBN
Kaguya	Sep. 2007	Moon	DARTS
Akatsuki	May 2010	Venus	DARTS PDS/Atmos.
Hayabusa2	Dec. 2012	Asteroid Ryugu	DARTS PDS/SBN
BepiColombo (MMO/MPO)	Oct. 2018 (Scheduled)	Mercury	DARTS PSA

For another example, there was an undefined coordinate system during a mission. The asteroid 25143 Itokawa was a target of the original Hayabusa mission. The definition of the coordinate was referred in Fujiwara et al. in 2006[4], and IAU officially approved the orientation parameters in 2007[5]. In this case, the precise values obtained by the original Hayabusa project affected the IAU definition.

Internal organization for PDS: For JAXA without many planetary probes, it is hard to organize specialized units of data archives for planetary explorations. Therefore, each project must build a data archive on a project basis. This structure works very well if the standard is not that complicated. Unfortu-natally, the planetary data archive is no longer simple already. The PDS is sophisticated for the long-term preservation and instrument team must understand abundant documents. It is required to prepare experts internally beyond missions to make a high-quality data archive continuously.

External support for PDS: While data standards are in place, securing resources to follow them is difficult. External support is essential to make archives high quality.

Indeed, developing PDS4 archives entailed many difficulties in Japan because JAXA has just acquired knowhow on the PDS3 standards. To confront this problem, the existence of the International Planetary Data Alliance (IPDA) was quite meaningful. There are

PDS4 related members in IPDA and JAXA was able to issue a question or a request to the PDS4 via IPDA.

Also, the contracts between NASA and JAXA, or ESA and JAXA are helpful. The memorandum of understanding (MOU) between agencies described the cooperation for data archives, and the framework beyond agencies was established.

Data status of Japanese planetary missions: JAXA opens scientific data obtained by spacecraft from the DARTS website[6].

Sakigake datasets were recently recreated to fill the coverage by the historical data recovery program in Japan. Originally, there was no data in DARTS, but newly archived datasets were released from DARTS in 2017. The format is a simple ASCII format, and still not in PDS3 format. We consider submitting this dataset to PDS Small Body Node(SBN) after PDS3 conversion.

The original Hayabusa was fully supported by MOU between NASA and JAXA to create archives in PDS3 format. Therefore, the data is currently released both from NASA's PDS website and JAXA's DARTS website.

The data archives of SELENE (Kaguya) mission was constructed in the original format similar to PDS3. They were converted from the original format to PDS3 over several years, and it was a very challenging task because the project was already finished and the cooperation of other agencies was minimum.

Archiving of data sets of Akatsuki (also known as Venus Climate Orbiter, VCO) are ongoing (in Nov. 2017). The SPICE kernels dataset had already been passed the review by PDS NAIF node, and is archived in DARTS and is mirrored by NAIF. The other datasets that mainly includes images acquired by the cam-eras onboard Akatsuki are currently under review by PDS Atmospheres Node.

Hayabusa2 is the first mission to implement PDS4 archives in the Japanese missions. To follow the new standards PDS4, the cooperation of NASA PDS and IPDA are essential. Training programs and talks to the actual developers of PDS4 are also available for the implementation.

Conclusion: Even with sophisticated design, it is difficult to implement. The data archive field of planetary exploration already requires experts. International cooperation is indispensable to enhance the quality of data archives.

References: [1] NASA The Planetary Data System (PDS) https://pds.nasa.gov,2017. [2] JPL NAIF https://naif.jpl.nasa.gov/naif/,2017. [3] LRO Project White paper Version 4 (2008) [4] Fujiwara A. (2006) *Science*, 312, 1330-1334. [5] Seidelmann P. K. (2007) *Celestial Mechanics and Dynamical Astr.*, 98-155. [6] JAXA/ISAS Data Archives and Transmission System (DARTS) http://darts.isas.jaxa.jp/, 2017.

THE ASTROMATERIALS X-RAY COMPUTED TOMOGRAPHY LABORATORY AT JOHNSON SPACE CENTER. R. A. Zeigler, E. H. Blumenfeld, P. Srinivasan, F. M. McCubbin, and C. A. Evans, Astromaterials Research and Exploration Division, NASA JSC, Mailcode XI, 2101 NASA Parkway, Houston, TX 77058. ryan.a.zeigler@nasa.gov.

Overview: The Astromaterials Acquisition and Curation Office at NASA's Johnson Space Center (hereafter JSC curation) is the past, present, and future home of all of NASA's astromaterials sample collections. JSC curation currently houses all or part of nine different sample collections. Our primary goals are to maintain the long-term integrity of the samples and ensure that the samples are distributed for scientific study in a fair, timely, and responsible manner, thus maximizing the return on each sample. Part of the curation process is planning for the future, thus we also perform fundamental research in advanced curation initiatives. Advanced Curation is tasked with developing procedures, technology, and data sets necessary for curating new types of sample collections, or getting new results from existing sample collections [1]. As part of these advanced curation efforts we are augmenting our analytical facilities. A micro X-ray computed tomography (micro-XCT) laboratory dedicated to the study of astromaterials came online within the JSC Curation office this summer, and we plan to add additional facilities that will enable non-destructive (or minimally-destructive) analyses of astromaterials in the near future (micro-XRF, confocal imaging Raman Spectroscopy). These facilities will be available to: (1) develop sample handling and storage techniques for future sample return missions, (2) be utilized by PET for future sample return missions, (3) be used for retroactive PET-style analyses of our existing collections, and (4) for periodic assessments of the existing sample collections. Here we describe the new micro-XCT system, as well as some of the ongoing or anticipated applications of the instrument, as well as the issues related to making the large volume of data available to the public.

Instrument: We have installed a Nikon XTH 320 micro-XCT system in JSC curation. It has four interchangeable X-ray sources: 180 kV nano focus transmission source, 225 kV reflection source with multi-metal target (Mo, W, Ag, Cu), a 225 kV rotating target (W) reflection source, and a 320 kV reflection source. The system also has a 16-bit, 400 mm^2 (2000 x 2000 pixel) CCD detector, as well as a heavy-duty stage that will accommodate large (up to 30 cm) and heavy (up to 100 kg) samples. The multiple sources, high-resolution detector, and large stage allow us the flexibility to analyze a wide range of sample sizes. The 180 kV trans-

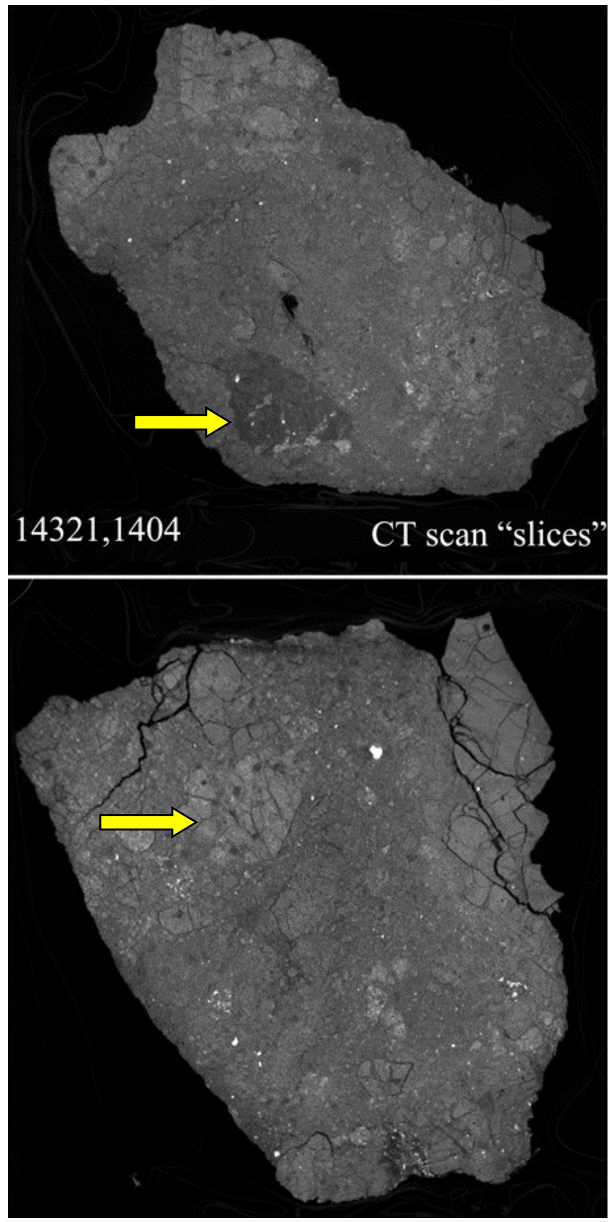


Figure 1: Slices of the micro-CT scan of sample 14321,1404. Brightness of the phases are proportional to x-ray attenuation (a measure of composition and density of the phase). Yellow arrows highlight interesting feldspathic (top) and mafic (bottom) clasts. Sample is ~ 6 cm in diameter.

mission source will allow for high resolution (submicron) scans on small samples (less than ~5 mm), whereas the 225 kV and 320 kV sources will allow scans of larger samples at resolutions on the order of 10s or 100s of microns per voxel depending on the sample size. The maximum size high-density rock sample that can be scanned has yet to be determined, but test scans on basalt samples >15 cm in diameter have been successful.

Discussion: High-intensity XCT scanners have been used to study astromaterials (and other geologic samples) for over 15 years [2-3], and the practice is becoming ever more prevalent. They have a wide range of scientific uses, including (but certainly not limited to) measuring porosity, determining the modal abundance and 3D distribution of phases inside samples, and identification of fabrics or strain patterns in samples. In addition to their use for research, XCT scans have increasingly been utilized as a part of the astromaterials curation process, beginning with meteorites [4-5], and more recently with the Apollo samples [6]. Their utility in curation lies in their ability to non-destructively map out the phases and voids within a sample. As an example, we have scanned several large Apollo polymict breccias, and we were able to identify and tentatively classify the lithologies in these clasts. The samples can then be subdivided, either through sawing or careful chipping, and those “new” clasts made available to scientists. In addition to the myriad curatorial uses, we have begun to use the XCT system as an integral part of coordinated analyses of astromaterials. An example of this is the mapping of textures within ungrouped achondrite NWA 11119 (Figure 2). The XCT scans were able to characterize the major phases with the meteorites, and identify areas of interest for additional higher resolution study (e.g., by TEM). The penetrative nature of the XCT scans allows for astromaterials samples to be analyzed within sealed low-density containers, preserving the pristinity of the samples. The XCT technique is not completely non-destructive, however. A recent study by [7] has shown that XCT scans of meteorites can alter the natural radiation dose of the sample. The number of techniques where this is applicable (e.g., thermo-luminescence) is limited, however. In the meantime, the percentage of any one sample that is studied by XCT will be limited to ensure that no irreparable damage is done to an entire sample. In addition to stand-alone XCT scans, we have an ongoing project to pair image-based 3D reconstructions of Apollo Lunar Samples correlated to 3D reconstruction of the same sample’s Micro-XCT data [8].

The data from XCT scans and related 3D imaging represents large volumes of data (many 10s of GB per scan) that will need to be curated for the long term and served to the public in a useful way; both are formibable challenges.

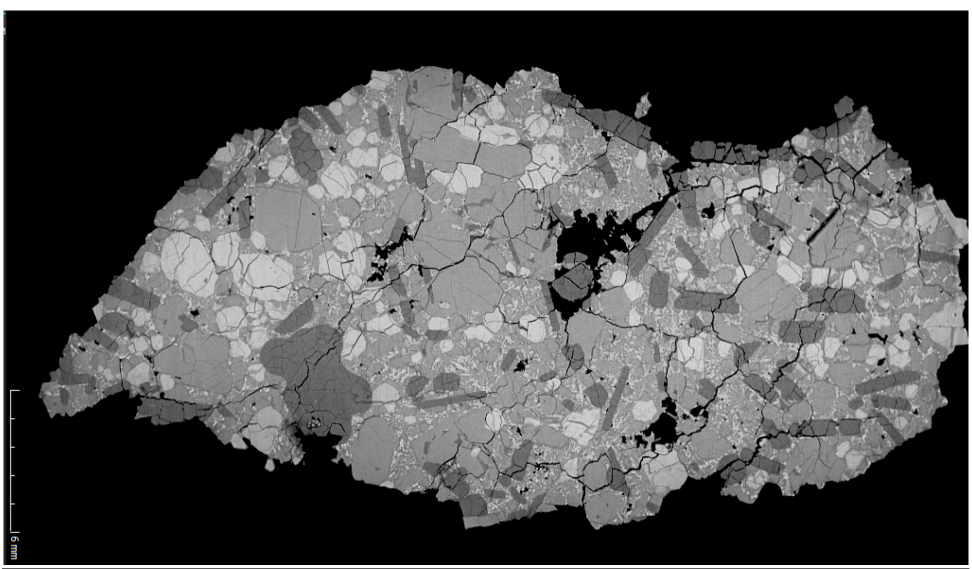


Figure 2: Single 2D slice of the micro-CT scan of ungrouped achondrite NWA 11119 showing the igneous textures and major mineralogical phases within the sample. Sample is ~ 3 cm long.

References [1] McCubbin et. al. (2016) 47th LPSC, abstract #2668. [2] McCoy et al. (2002) EPSL 246, 102-108. [3] Zolensky et al. (2014) MAPS 49, 1997-2016. [4] Almeida et al. (2014), MAPS 77 abstract #5033. [5] Smith C. L. (2013) MAPS 76, abstract #5323. [6] Zeigler (2014) 45th LPSC, abstract #2665. [7] Sears, et al. (2016) MAPS 51.4, 833-838. [8] Blumenfeld et al. (2016) AGU abstract # ED42B-05.

Lunar Circular Structure Classification from Chang ’e 2 High Resolution Lunar Images with Convolutional Neural Network. X. G. Zeng[1] , J. J. Liu[1], Wei. Zuo[1], W. L. Chen[1]and Y. X. Liu[1], [1] Key Laboratory of Lunar and Deep Space Exploration, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China,zengxg@nao.cas.cn,liujj@nao.cas.cn,zuowei@nao.cas.cn,chenwl@nao.cas.cn,liuyx@nao.cas.cn

Introduction: Circular structures are widely distributed around the lunar surface, the most typical of them could be lunar impact crater, lunar dome et.al[1]. By the identification and analysis of these structures, the geological evolution process and a lot of other information about the Moon could be revealed. Most researchers have tried to identify these structures by recognition of their circular morphology from image data or topography data. With the image or topographic data, traditional edge extraction algorithms(such as SVM), and image data classification algorithm(such as MLE) are frequently used. With these method, most of the circular structures in the lunar surface could be located and a lot of structure database are originated from them. However, there still exist some drawback among these method, such as, the accuracy rate of identified result from these method are unstable which could upto 60%-80%, which means there are still many errors and missing among the result, however, it is very hard the improve the accuary rate with these traditional methods.

The great success of image calssfication with deep learning has attracted the attention of planetary science community, and many scientists are also trying to use deep learning method for image understanding in planetary science[2][3]. And, like the traditional image classification method, there are supervised classification and unsupervised classification in deep learning method, and Convolutional Neural Network(CNNs) is a supervised classification model. A typical CNN is comprised of one or more convolutional layers (often with a subsampling step) and then followed by one or more fully connected layers as in a standard multilayer neural network. The architecture of a CNN is designed to take advantage of the 2D structure of an input image. This is achieved with local connections and tied weights followed by some form of pooling which results in translation invariant features, which means it is good for image recognition. Another advantage of CNNs is that they are easier to train and have many fewer parameters than fully connected networks with the same number of hidden units. So, in this approach, we are trying to use the CNN method to classfy the lunar circular structures from the lunar images.

Method: Since it is difficult to build a CNN totally from the fundamental, while there are servral commercial deep learning framework such as Caffe,Theano,Torch, and TensorFlow, in this approach, we are going to build the CNN with the help of the Caffe Framework.

Preparation with Caffe. Caffe is a deep learning framework made with expression, speed, and modularity in mind[4]. As the request of this framework, we need to prepare the supervised sample images for cicular structures, non-circular structures,label these two kinds data with label like ‘1’ and ‘2’(see Figure 1 and Figure 2), and divide the data into train dataset and value dataset, both in the specific data format required by the system; after that, we also need to select a CNN net model (such as AlexNet), the parameters of the net model needed to be modified; then we need to train the model with train dataset and export the model into a model file; and finally with the model file, we could classify the value dataset and get the classification result.

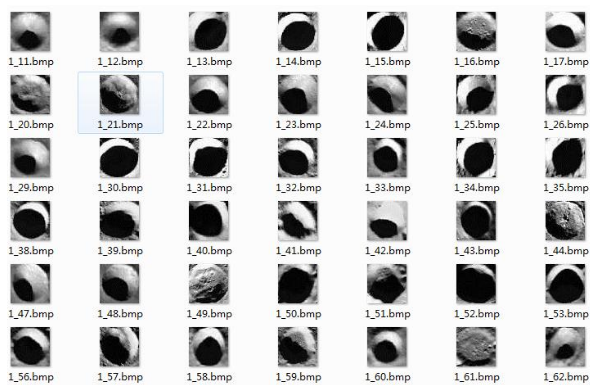


Figure 1 Circular sample images

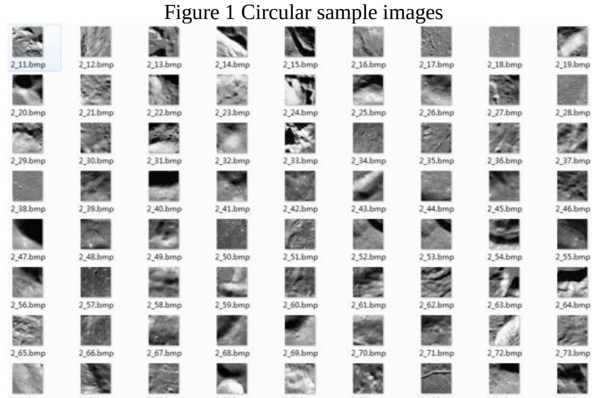


Figure 2 Non-Circular sample images

About the lunar image data. The source data of the lunar samele images are processed from the Chang’e 2 data, the resolution of the original data is 7m. We calibrated and divided each image into a single image with the size of 48pixel*48pixel.

Discussion: The study of this work is still under work, and the initial result of the work shows that the model of CNN needs a large amount of sample images, both for circular and non circular sample images, and the training rate is not so good at this moment, we have analysed the sample images and trying to use different net model, ajust the parameters to improve the accuracy of the model.

References:

[1] Pike RJ (1977). "Size-dependence in the shape of fresh impact craters on the moon.". Impact and explosion cratering: Planetary and terrestrial implications; Proceedings of the Symposium on Planetary Cratering Mechanics, Flagstaff, Ariz., September 13-17, 1976. New York: Pergamon Press. pp. 489–509.

[2] Jasper L E Z, Xaypraseuth P. Data production on past and future NASA missions[C]//Aerospace Conference, 2017 IEEE. IEEE, 2017: 1-11..

[3] Cupani G, Cristiani S, D'Odorico V, et al. E-ELT HIRES the high resolution spectrograph for the E-ELT: integrated data flow system[C]//SPIE Astronomical Telescopes+ Instrumentation. International Society for Optics and Photonics, 2016: 991023-991023-10.

[4] Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional architecture for fast feature embedding[C]//Proceedings of the 22nd ACM international conference on Multimedia. ACM, 2014: 675-678.

MATISSE 2.0: new ideas to support planetary sciences. A. Zinzi[1,2], A. Longobardo[3], M. Giardino[1], S. Ivanovski[3], M.T. Capria[3], E. Palomba[3], [1]SSDC-ASI (angelo.zinzi@ssdc.asi.it – Via del Politecnico, snc – 00133 – Rome, Italy), [2]INAF-OAR, [3]INAF-IAPS.

Introduction: Five years after the beginning of the project, MATISSE (Multi-purpose Advanced Tool for Instruments for the Solar System Exploration – https://tools.asdc.asi.it/matisse.jsp, [1]) now ingests observations in different formats and from different space missions and targets.

Its main aim was to provide the community with a tool able to make easier and enhance the scientific production of a single instrument: this goal has been pursued by making the data searchable with a metadata driven database, introducing the possibility of computing high-order products (i.e., mosaics and ratios) and visualize them in advanced forms.

The tool presently allows to search, analyze and visualize data regarding comet 67P Churyumov-Gerasimenko, asteroid 4 Vesta, dwarf panet 1 Ceres, the Moon and the planets Mercury, Venus and Mars.

In particular for these three latter objects new approaches have been developed, allowing the data not to be physically located at ASI Space Science Data Center (SSDC), but to be queried over remote servers (e.g., [2]).

The expansion of MATISSE to objects larger than asteroids and comets, for which the tool had been specifically designed at the beginning, now requires a robust update, so that valuable online visualizations can be displayed for all the targets.

Present-day MATISSE: MATISSE currently uses a series of programming languages in order to start the pipeline, read the observation files, perform computations and generate the output objects.

These are composed by two static PNG maps, a 3D online visualization (only for smaller targets), and downloadable files, such as 2D projections in ENVI, GeoTIFF and FITS file and a 3D projection as a VTP file.

This architecture fullfils the original requirements of searching for observations inside a local or remote database, obtaining mosaics and ratios merging together single observations and displaying the results in both 2D and 3D format for online and desktop usage.

However it also introduce several limitations that, with the time, need to be overcome.

MATISSE 2.0: The next version of MATISSE will be a brand-new one, so that, without wiping out the succesfull modular structure of the current version, a number of newly introduced solution can allow to overcome existing limitations.

In particular, the pipeline will be completely written in Python, except where software readers are available only in different languages (e.g. LecturePDS, written in IDL/GDL [3]).

Hence it will be possible to adopt solutions able to speed up the computations, for example by avoiding to write intermediate files to be exchanged between scripts in different languages, or by using parallelization for saving machine-time when possibile.

Another main difference will be in the output page, where the 3D visualization will be performed using the same X3DOM javascript presently used, but adopting binary files instead of ASCII ones. This will add the possibility of displaying inside the browser also high-resolution versions of large planetary Digital Elevation Models.

Furthermore, the static maps will be replaced by FITS projected images visualized by means of the JS9 suite. In this way it will be possible for the user to change zoom, palettes and scales in order to fit the output requirements.

By also changing the graphic interface, making it possible to improve responsivity, user support and usability of the tool, will be one of the target of the newly designed version of the tool.

Conclusions: The totally new version of ASI SSDC MATISSE tool, planned to be completed for the end of 2018, will add a number of valuable features to the tool.

The new version will allow planetary sciences community a better usage of the tool, that could be used for enhanced scientific purposes, without the limitations arssen at the completion of its first version.

References:

[1] Zinzi A. et al. (2016), Astronomy and Computing, 15, 16-28

[2] http://openplanetary.co/blog/tool/matisse2ps.html

[3] Erard S. et al. (2013), http://vopariseuroplanet.obspm.fr/utilities/Virtispds_3.1a.pdf

PACKMAN-Net: A Distributed, Open-Access and Scalable, Network of User-Friendly Space Weather Stations.

M.-P. Zorzano[1,2], J. Martín-Torres [1,3], T. Mathanlal[1], A. Vakkada Ramachandran[1] and J.-A. Ramirez-Luque[1]. [1]Atmospheric Science Group, Department of Computer Sciences, Electrical and Space Engineering, Luleå University of Technology, 971 87 Luleå, Sweden, [2] Centro de Astrobiología (INTA-CSIC), Torrejón de Ardoz, Madrid, Spain. [3] Instituto Andaluz de Ciencias de la Tierra (CSIC-UGR), Armilla, Granada, Spain.

Introduction: The Earth's atmosphere is continuously bombarded by energetic charged particles from space. To date, there is a missing gap of information regarding the amount, energy, time variability, and type of space radiation that reaches the lower layers of the atmosphere, as well as on its geographic and altitude distribution. This information has implications for space weather, and potential impact on infrastructures and climate. To generate a long-time record of space radiation on Earth we designed an open source, autonomous instrument, called PACKMAN (PArticle Counter k-index Magnetic ANomaly), with Commercial Off The Shelf (COTS) components, that can monitor continuously a set of critical variables at multiple latitudes and heights in the atmosphere (in balloons) [1]. After an initial deployment and testing phase, PACKMAN has demonstrated its operability at different latitudes and atmospheric heights (see *Initial deployment* section below). The purpose of the PACKMAN-Net phase of the project is twofold: 1) to design an architecture that allows for scalability with multiple instrumental nodes of equivalent format that adhere to the PDS4 standard; and 2) to ensure an efficient management (discovery, access, retrieval and analysis) of the information provided by a network of PACKMAN instruments distributed over the world by merging information and preparing data mining tools both for science, education and outreach, and space-weather alert warning systems. It aims to allow global electronic access and to provide global monitorization of space weather on community-based ground observatories as well as in other atmospheric platforms such as balloons.

This is an example of instrumentation data in our planet which needs a dedicated scalable data archiving and processing architecture that adheres to the PDS4 standard. Furthermore, this project may be used to benchmark the design of archiving, scalable, networks of future planetary instrumentation observations of the Moon or Mars.

Instrument network specifications: The purpose of this work is to demonstrate the operability utility of a network of small-sized detectors of the PACKMAN instrument, operated simultaneously to provide real time cosmic ray and solar activity monitoring, covering ground based observations over the entire planet. A critical and complementary observation is the one from atmospheric vertical soundings, using the same hardware, and measuring simultaneously ideally in a long duration (several days) circumpolar campaign, but also on shorter stratospheric and tropospheric sounding balloons campaigns.

PACKMAN is an autonomous instrument that can be deployed at any location and send the data automatically through wireless communications. The PACKMAN-G (ground) is adapted for surface monitoring, including outdoors remote operation, to provide simultaneous records at multiple latitudes (and longitudes) with different geomagnetic fields and at different heights with different total air column (pressure) and weather phenomena. The PACKMAN-B (balloon) is a TRL8 qualified flight Instrument used in balloon campaigns for tropospheric and stratospheric sounding. PACKMAN-B has been flown in two stratospheric balloons from Córdoba, Spain with Zero2Infinity and from Esrange Space Center, Kiruna, Sweden in collaboration with Swedish Space Corporation (SSC).

Initial deployment:

Through the initial deployment and testing phase, PACKMAN has demonstrated its operability at multiple latitudes and atmospheric heights:

1. Space campus LTU, Kiruna, Sweden (67.84°N, 20.41°E, 390 m)
2. LTU Main campus, Luleå, Sweden (65.62°N, 22.14°E, 15 m)
3. Boulby Mine, Cleaveland, United Kingdom (54.56°N, 0.82°W, 93 m and -1.1 km)
4. University of Edinburgh, United Kingdom (55.94°N, 3.19°W, 98 m)
5. Cordoba airport, Córdoba, Spain (37.84°N, 4.84°W, 90 m to 27 km)
6. Esrange Space Center, Kiruna, Sweden (67.88°N, 21.12°E, 328 m to 27 km)

PACKMAN is a small, robust, light and scalable instrument that monitors with two Geiger counters gamma, beta, alpha radiation and muons. This instrument includes environmental sensors to monitor pressure, temperature, relative humidity, and magnetic perturbations (with three fluxgate magnetometers in three perpendicular axes) and includes data archiving, GPS and communication capabilities.

We illustrate through this initial network, that PACKMAN can unequivocally detect the onset of a solar storm and can provide meaningful records of particle counts, synchronous high energy particle detections, pressure and/or density anti-correlations, dependence on local geomagnetic field and elevation. We demonstrate the successful operation of the flight-model of PACKMAN instrument from the surface to the stratosphere, in a space operational environment. Through these testing campaigns, the technology has been proven to work in its final form, both for ground and space operation, and under nominal operation conditions.

Data archiving and access specifications: At present date there are 4 PACKMAN nodes operating continuously at different latitudes, and 2 extra PACKMAN nodes will be installed in Granada, Spain, and Misasa, Japan durint the spring of 2018. Besides the natural expansion of the network through collaborations with research institutions, in the future, when the instructions for construction are released, it is expected that more users may adhere to the network by building their own instrument and uploading the data. This will allow to make a scalable network and also bridge the gap between society and research, adding new stake holders such as teaching institutions (high schools, universities), or industry and infrastructures representatives. PACKMAN data will be formatted in PDS4 standard, in such a way that every instrument will be an instrument node with their own structure. New instruments can be integrated in the server creating new nodes and uploading the data over time to the same server. To be able to upload data, credentials will be required and anyone with a PACKMAN instrument can request them in PACKMAN dedicated website that will be available in a near future.

To have consistency in the products that we create, we will provide a software to save the data in the same format (for instance using the same time stamp label format and same order of variables, and file size or cadence sequence) and to upload the data to the system. Alternatively, and for users who only want to format their data, we will provide the necessary software to save the data in the correct format, this can be used for testing the instrument or to have a PACKMAN unit working as a standalone product.

The website will contain PACKMAN general information, documentation, software and registration procedure to access / upload the data. Furthermore a set of online tools and data mining codes will alow to analyse the data from PACKMAN units distributed over the world by comparing the instrument observations and providing new products for science, education and outreach, and space-weather alert warning systems. The present design considers the possibility to merge the data with other instruments (including space weather observations from orbit) and to increase the scope of research of the PACKMAN network by including extra sensors.

The observations acquired by PACKMAN will be used to provide open access, real time information, for: 1) education and public awareness of space weather phenomena; 2) to compare with Earth climate observations; 3) to provide real-time information of space weather variability for potential damage to infrastructures (telecommunications, power generation facilities, aviation, transport, etc.); 4) to monitor natural radiation sources at multiple environments; 5) to monitor the variability of the Pfotzer maximum height during different stages of solar activity and seasons and 6) finally, this project may serve as a reference for future scalable networks where multiple instruments are deployed at different sites or conditions and with different initiation times, and where the informational value increases by adhering to a common PDS4 format and analysing the data in a concurrent way.

References:

[1] Zorzano M.-P. et al (2017). *The PACKMAN radiation and environmental instrument for space weather studies.* PAC Proceedings ESA. 23rd ESA Symposium on European Balloon and Rocket Programmes and related Research – 11th -15th June 2017, Visby, Sweden.